

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 January 2001 (18.01.2001)

PCT

(10) International Publication Number  
**WO 01/04270 A1**

- (51) International Patent Classification<sup>7</sup>: **C12N 5/10**, 15/12, 15/63, 15/64, C07K 14/47, 14/705, 14/715, G01N 33/53, 33/567
- (21) International Application Number: **PCT/US00/19094**
- (22) International Filing Date: **13 July 2000 (13.07.2000)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:  
**60/143,435** **13 July 1999 (13.07.1999)** **US**
- (71) Applicant (*for all designated States except US*): **GENAIS-SANCE PHARMACEUTICALS, INC.** [US/US]; Five Science Park, New Haven, CT 06511 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **CHEW, Anne** [US/US]; 837 Orange Street, Apartment 2, New Haven, CT 06511 (US). **DENTON, R., Rex** [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). **DUDA, Amy** [US/US]; 93 Marvel Road, New Haven, CT 06515 (US). **NAND-ABALAN, Krishnan** [IN/US]; 228 Village Pond Road, Guilford, CT 06437 (US). **STEPHENS, Joel, Claiborne** [US/US]; 46 Crabapple Lane, Guilford, CT 06437 (US). **WINDEMUTH, Andreas, K.** [US/US]; 91 Center Road, Woodbridge, CT 06524 (US).
- (74) Agent: **HENDERSON, Melodie, W.**; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).
- (81) Designated States (*national*): **AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.**
- (84) Designated States (*regional*): **ARIPO** patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), **Eurasian** patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), **European** patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), **OAPI** patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
- *With international search report.*
  - *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/04270 A1

(54) Title: **DRUG TARGET ISOGENES: POLYMORPHISMS IN THE INTERLEUKIN 4 RECEPTOR ALPHA GENE**

(57) Abstract: Polynucleotides comprising one or more of 38 novel single nucleotide polymorphisms in the human Interleukin 4 Receptor Alpha (IL4R $\alpha$ ) gene are described. Compositions and methods for detecting one or more of these polymorphisms are also disclosed. In addition, various genotypes and haplotypes for IL4R $\alpha$  gene that exist in the population are described.

## DRUG TARGET ISOGENES:

## POLYMORPHISMS IN THE INTERLEUKIN 4 RECEPTOR ALPHA GENE

## RELATED APPLICATIONS

This application is a continuation-in-part of, and claims priority to, U.S. Provisional Application  
Serial No. 60/143,435 filed July 13, 1999.

## FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically important proteins. In particular, this invention provides genetic variants of the human Interleukin 4 Receptor Alpha (IL4R $\alpha$ ) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

## BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a drug that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by a compound's activity at non-intended targets.

What this approach fails to consider, however, is that natural variability exists in any and every population with respect to a particular protein. A target protein currently used to screen drugs typically is expressed by a gene cloned from an individual who was arbitrarily selected. However, the nucleotide sequence of a particular gene may vary tremendously among individuals. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a target protein may be manifested as significant variation in expression of or in the structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in treatment of individuals with drugs whose design is based upon a single representative example of the target. For example, it is well-established that some classes of drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. In addition, variable information on the biological function or effects of a particular protein may be due to different scientists unknowingly studying different isoforms of the gene encoding the protein. Thus, information on the type and frequency of genomic variation that exists for pharmaceutically important proteins would be useful.

The organization of single nucleotide variations (polymorphisms) in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a haplotype. Each haplotype therefore contains significantly more information than individual unorganized polymorphisms. Haplotypes provide an accurate measurement of the genomic variation in the two chromosomes of an individual.

It is well-established that many diseases are associated with specific variations in gene sequences. However while there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* 63:595-612; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* 161: 469-74). In addition, the marker may be predictive in some populations, but not in other populations (Clark AG et al. 1998 *supra*). In these instances, a haplotype will provide a superior genetic marker for the phenotype (Clark AG et al. 1998 *supra*; Ulbrecht M et al. 2000, *supra*; Ruaño G & Stephens JC *Gen Eng News* 19 (21), December 1999).

Analysis of the association between each observed haplotype and a particular phenotype permits ranking of each haplotype by its statistical power of prediction for the phenotype. Haplotypes found to be strongly associated with the phenotype can then have that positive association confirmed by alternative methods to minimize false positives. For a gene suspected to be associated with a particular phenotype, if no observed haplotypes for that gene show association with the phenotype of interest, then it may be inferred that variation in the gene has little, if any, involvement with that phenotype (Ruaño & Stephens 1999, *supra*). Thus, information on the observed haplotypes and their frequency of occurrence in various population groups will be useful in a variety of research and clinical applications.

One possible drug target for the treatment of allergies, asthma, and other immune responses is the Interleukin 4 Receptor Alpha (IL4R $\alpha$ ) gene or its encoded product. IL4R is a transmembrane complex composed of two different protein subunits, a 140-kDa high affinity binding subunit named interleukin-4 receptor  $\alpha$  (IL-4R $\alpha$ ; also known as CD124 antigen) and either a gamma-c subunit, which is present in several cytokine receptors, or an interleukin-13 receptor 1 (IL-13R1) subunit. Both subunits of the IL-4R are required to bind interleukin-4 (IL-4) and to mediate its transcription-activating effects through the tyrosine kinases, Jak1 and Jak3. Upon binding of IL-4 to the IL-4R, Jak 1 and Jak 3 phosphorylate the IL-4R $\alpha$  subunit, creating binding sites in the cytoplasmic domain for many other proteins, including SOS, Stat-6, c-fes, and *src* homology phosphatase 1 (SHP-1). Activated Jak proteins also phosphorylate Stat-proteins, which travel into the nucleus and function as transcription factors. Other IL-4 signal pathways exist, but are less well characterized.

IL-4, one of the most important cytokines involved in the allergic response, is produced when cells from the immune system, in particular T cells, are activated in response to an allergen. Regulation of the immune response involves Helper T-cells that differentiate into two subtypes, Th1 and Th2. Th1 cells express interferon-gamma and interleukin-2 (IL-2), and mediate a cell-based immunity, where macrophages and neutrophils are prominently involved. Th1 cells also direct the IgE-producing B cells, as well as mast cells, basophils and eosinophils. Th2 cells produce IL-4, IL-5, IL-6, IL-10 and IL-13. Each Th cell subtype represses the other, so the immune system is forced into differentiation into either a Th1 or Th2 response against an external allergic challenge. In many instances an aberration of this response can render a pathological state such as a Th2 response against ragweed.

IL-4 induces in B cells the synthesis of IgE type antibodies that recognize specific allergens. IgE binds to receptors on mast cells and basophils and mediates the early humoral (sub-chronic) response on the B-side of the immune system. If an allergen binds mast cell-attached IgE, the mast cell releases mediators like histamine, and the eicosanoid leukotrienes and prostaglandins products, some of which cause the familiar symptoms of an acute allergic reaction: swelling, itching, mucous, and reddening of the skin. Later in this process eosinophils and other inflammatory cells migrate to the site of inflammation. This later phase is important in asthma, because the eosinophils may instigate a more chronic inflammation which can adversely scar lung tissue. IL-4 is at least partly responsible for recruiting eosinophils, because it induces synthesis of specific adhesion molecules on the capillary endothelium, and stimulates expression of IL-5 and eotaxin. IL-5 leads to the development of a large number of eosinophils from precursor cells in the bone marrow, and eotaxin stimulates their migration into the lung tissue.

It has been proposed that inhibition of IL-4 activity would disarm the Th2 component of the immune system. This would then allow the immune system to develop a natural tolerance towards common allergens without the full acute response to the challenge. In this way, tolerance may be induced in many patients, similar to what is sometimes achieved with hyposensitization shots for allergy patients. Thus, substances that inhibit IL-4 production and/or its binding to the IL-4 receptor, may improve the therapy of allergies and asthma.

The gene for IL-4R $\alpha$  has eleven exons encoding an 825 amino acid protein and spans over 24 kb of the short arm of chromosome 16 (16p12.1) (Pritchard et al., *Genomics* 10:801, 1991; GenBank Accession No. AC004525). A reference sequence for IL4R- $\alpha$  gene, which corresponds to the reverse complement of nucleotides 100020-71331 in the GenBank Accession No. AC004525, is shown in Fig. 1 (SEQ ID NO:1). Reference sequences for IL-4R $\alpha$  mRNA (GenBank Accession No. NM\_000418) and the encoded IL4R $\alpha$  precursor protein (GenBank Accession No. P24394) are shown in Figs. 2 and 3, respectively (SEQ ID NOS:2 and 3). Significant features reported for the IL-4R $\alpha$  precursor include: a signal peptide located between a.a. 1 and 25; an extracellular domain between a.a. 26 and 232; disulfide bonds between a.a. 34 and 44 and between a.a. 74 and 86; glycosylation sites at amino acids 53, 98, 128, 134, 176, and 209; a transmembrane region between a.a. 233 and 256; and a cytoplasmic domain between a.a. 257 and 825.

Recently, several studies have suggested that genetic polymorphisms in the IL4-R $\alpha$  gene are associated with genetic predisposition to atopy and/or elevated serum IgE. Mitsuyasu et al., reported that polymorphisms at codons 75 and 576 affect IL-4R function (*Nat. Genet.* 19:119-120, 1998). The IL-4R $\alpha$  allele with isoleucine at amino acid position 75 (Ile75) in the extracellular domain is more responsive to IL-4 than the allele with valine at that position (Val75) and is associated with atopic asthma but not with non-atopic asthma (Mitsuyasu et al., *supra*). Also, the allele with arginine at position 576 (Arg576) in the cytoplasmic domain exhibits higher receptor activity than the glutamine allele (Glu576) due to reduced



binding by the Arg576 allele of a negative regulatory molecule, *src* homology phosphatase 1 (Imani et al., *J. Biol Chem.* 272:7927-7931, 1997). The Arg576 allele has a higher frequency in patients with allergic inflammatory disorders, including atopy (Khurana Hershey et al., *New Eng. J. Med.* 337:1720-1725, 1997). In a recent study, Hershey et al.; (WO 00/34789) reported that the Arg576 allele is significantly associated with asthma. Studies showed that patients who were homozygous for this allele had about a 9-fold higher risk towards asthma and that two copies of Arg576 are associated in an increase in asthma prevalence and severity. Kanemitsu et al. recently reported that the presence of either the Ile75 or Arg576 variant is significantly associated with susceptibility for developing systemic lupus erythematosus (SLE), a Th2-dominant systemic autoimmune disorder (*Arthritis Rheum.* 42:1298-1300). Another variant IL-4R $\alpha$  allele that is reportedly associated with atopy susceptibility has proline rather than serine at position 503 (Kruse et al., *Immunol.* 96:365-371, 1999). Other IL-4R $\alpha$  gene polymorphisms leading to amino acid changes in the cytoplasmic domain of the protein product have been identified at codons 400 (E400A), 431 (C431R) and 786 (S786P) (Kruse et al., *supra*; Deichmann et al., *Biochem. Biophys. Res. Commun.* 231:696-697, 1997). A polymorphism of guanine or adenine at a position corresponding to nucleotide 55328 in Figure 1 has also been reported as well as a polymorphism of cytosine or thymine at a position corresponding to nucleotide 55430 (Buetow et al., 1999, *Nat Genet.* 21:323-5).

Because of the potential for polymorphisms in the IL4R $\alpha$  gene to affect the expression and function of the encoded protein, it would be useful to determine whether additional polymorphisms exist in the IL4R $\alpha$  gene, as well as how such polymorphisms are combined in different copies of the gene.

Such information would be useful for studying the biological function of IL4R $\alpha$  as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

#### SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 38 novel polymorphic sites in the IL4R $\alpha$  gene. These polymorphic sites (PS) correspond to the following nucleotide positions in the reverse complement of the indicated GenBank Accession Number: 32884 (PS1), 32903 (PS2), 32961 (PS3), 33135 (PS4), 35763 (PS6), 35770 (PS7), 35817 (PS8), 35905 (PS9), 35944 (PS10), 35958 (PS11), 37330 (PS12), 37473 (PS13), 37586 (PS14), 37591 (PS15), 37604 (PS16), 37644 (PS17), 37678 (PS18), 43446 (PS19), 43703 (PS20), 53008 (PS21), 53099 (PS22), 53153 (PS23), 53456 (PS25), 53507 (PS27), 53513 (PS28), 53915 (PS30), 53949 (PS32), 54237 (PS33), 54468 (PS34), 54611 (PS35), 54698 (PS36), 54700 (PS37), 54741 (PS38), 54780 (PS39), 55083 (PS40), 55142 (PS41), 55539 (PS44) and 55758 (PS45) in AC004525. The polymorphisms at these sites are adenine or guanine at PS1, cytosine or thymine at PS2, guanine or thymine at PS3, guanine or cytosine at PS4, cytosine or thymine at PS6, guanine or adenine at PS7, thymine or cytosine at PS8, cytosine or thymine at PS9, cytosine or thymine at PS10, guanine or adenine at PS11, guanine or adenine at PS12, cytosine or thymine at PS13, cytosine or thymine at PS14, guanine or adenine at PS15, adenine or thymine at PS16, cytosine or adenine at PS17, cytosine or

thymine at PS18, guanine or adenine at PS19, thymine or cytosine at PS20, adenine or cytosine at PS21, cytosine or thymine at PS22, thymine or cytosine at PS23, guanine or thymine at PS25, cytosine or thymine at PS27, thymine or cytosine at PS28, cytosine or thymine at PS30, guanine or adenine at PS32, cytosine or thymine at PS33, thymine or guanine at PS34, thymine or cytosine at PS35, thymine or cytosine at PS36, thymine or cytosine at PS37, cytosine or thymine at PS38, cytosine or guanine at PS39, adenine or guanine at PS40, guanine or adenine at PS41, cytosine or thymine at PS44 and guanine or adenine at PS45. In addition, the inventors have determined the identity of the alternative nucleotides present at these sites, as well as at the previously identified sites at nucleotides 35749 (PS5), 53413 (PS24), 53505 (PS26), 53721 (PS29), 53941 (PS31), 55328 (PS42), and 55430 (PS43). It is believed that IL4R $\alpha$ -encoding polynucleotides containing one or more of the novel polymorphic sites reported herein will be useful in studying the expression and biological function of IL4R $\alpha$ , as well as in developing drugs targeting this protein. In addition, information on the combinations of polymorphisms in the IL4R $\alpha$  gene may have diagnostic and forensic applications.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IL4R $\alpha$  gene or a fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of guanine at PS1, thymine at PS2, thymine at PS3, cytosine at PS4, thymine at PS6, adenine at PS7, cytosine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, adenine at PS12, thymine at PS13, thymine at PS14, adenine at PS15, thymine at PS16, adenine at PS17, thymine at PS18, adenine at PS19, cytosine at PS20, cytosine at PS21, thymine at PS22, cytosine at PS23, thymine at PS25, thymine at PS27, cytosine at PS28, thymine at PS30, adenine at PS32, thymine at PS33, guanine at PS34, cytosine at PS35, cytosine at PS36, cytosine at PS37, thymine at PS38, guanine at PS39, guanine at PS40, adenine at PS41, thymine at PS44, and adenine at PS45. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of guanine at PS5, cytosine at PS24, cytosine at PS26, cytosine at PS29, guanine at PS31, adenine at PS42, and thymine at PS43. A particularly preferred polymorphic variant is a naturally-occurring isoform (also referred to herein as an "isogene") of the IL4R $\alpha$  gene. An IL4R $\alpha$  isogene of the invention comprises adenine or guanine at PS1, cytosine or thymine at PS2, guanine or thymine at PS3, guanine or cytosine at PS4, cytosine or thymine at PS6, guanine or adenine at PS7, thymine or cytosine at PS8, cytosine or thymine at PS9, cytosine or thymine at PS10, guanine or adenine at PS11, guanine or adenine at PS12, cytosine or thymine at PS13, cytosine or thymine at PS14, guanine or adenine at PS15, adenine or thymine at PS16, cytosine or adenine at PS17, cytosine or thymine at PS18, guanine or adenine at PS19, thymine or cytosine at PS20, adenine or cytosine at PS21, cytosine or thymine at PS22, thymine or cytosine at PS23, guanine or thymine at PS25, cytosine or thymine at PS27, thymine or cytosine at PS28, cytosine or thymine at PS30, guanine or adenine at PS32, cytosine or thymine at PS33, thymine or guanine at PS34, thymine or cytosine at PS35, thymine or cytosine at PS36, thymine or cytosine at PS37, cytosine or thymine at PS38, cytosine or

guanine at PS39, adenine or guanine at PS40, guanine or adenine at PS41, cytosine or thymine at PS44 and guanine or adenine at PS45. The invention also provides a collection of IL4R $\alpha$  isogenes, referred to herein as an IL4R $\alpha$  genome anthology.

5 An IL4R $\alpha$  isogene may be defined by the combination and order of these polymorphisms in the isogene, which is referred to herein as an IL4R $\alpha$  haplotype. Thus, the invention also provides data on the number of different IL4R $\alpha$  haplotypes found in the above four population groups. This haplotype data is useful in methods for deriving an IL4R $\alpha$  haplotype from an individual's genotype for the IL4R $\alpha$  gene and for determining an association between an IL4R $\alpha$  haplotype and a particular trait.

10 In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for an IL4R $\alpha$  cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig. 2) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 237, adenine at a position corresponding to nucleotide 244, cytosine at a position corresponding to nucleotide 291, thymine at a position corresponding to nucleotide 501, adenine at a position corresponding to nucleotide 554, cytosine at a position corresponding to nucleotide 939, thymine at a position corresponding to nucleotide 1242, thymine at a position corresponding to nucleotide 1293, cytosine at a position corresponding to nucleotide 1299, thymine at a position corresponding to nucleotide 1701, adenine at a position corresponding to nucleotide 1735, thymine at a position corresponding to nucleotide 2023, guanine at a position corresponding to nucleotide 2254 and cytosine at a position corresponding to nucleotide 2397. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of guanine at a position corresponding to 223, cytosine at a position corresponding to nucleotide 1199, cytosine at a position corresponding to 1291, cytosine at a position corresponding to nucleotide 1507 and guanine at a position corresponding to 1737.

20 Polynucleotides complementary to these IL4R $\alpha$  genomic and cDNA variants are also provided by the invention.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express IL4R $\alpha$  for protein structure analysis and drug binding studies.

30 In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the IL4R $\alpha$  protein. The reference amino acid sequence comprises SEQ ID NO:3 (Fig. 3) and the polymorphic variant comprises at least one variant amino acid selected from the group consisting of threonine at a position corresponding to amino acid 82, histidine at a position corresponding to amino acid 185, isoleucine at a position corresponding to amino acid 579, serine at a position corresponding to amino acid 675, and alanine at a position corresponding to amino acid 752. In some embodiments, the polymorphic variant also comprises at least one variant amino acid

selected from the group consisting of valine at a position corresponding to amino acid 75, alanine at a position corresponding to amino acid 400, arginine at a position corresponding to amino acid 431, proline at a position corresponding to amino acid 503, and arginine at a position corresponding to amino acid 576. A polymorphic variant of IL4R $\alpha$  is useful in studying the effect of the variation on the biological activity of IL4R $\alpha$  as well as studying the binding affinity of candidate drugs targeting IL4R $\alpha$  for the treatment of allergies, asthma, and other immune responses.

The present invention also provides antibodies that recognize and bind to the above polymorphic IL4R $\alpha$  protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

In other embodiments, the invention provides methods, compositions, and kits for haplotyping and/or genotyping the IL4R $\alpha$  gene in an individual. The methods involve identifying the nucleotide or nucleotide pair present at one or more polymorphic sites selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45 in one or both copies of the IL4R $\alpha$  gene from the individual. The compositions contain oligonucleotide probes and primers designed to specifically hybridize to one or more target regions containing, or that are adjacent to, a polymorphic site. The methods and compositions for establishing the genotype or haplotype of an individual at the novel polymorphic sites described herein are useful for studying the effect of the polymorphisms in the etiology of diseases affected by the expression and function of the IL4R $\alpha$  protein, studying the efficacy of drugs targeting IL4R $\alpha$ , predicting individual susceptibility to diseases affected by the expression and function of the IL4R $\alpha$  protein and predicting individual responsiveness to drugs targeting IL4R $\alpha$ .

In yet another embodiment, the invention provides a method for identifying an association between a genotype or haplotype and a trait. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. Such methods have applicability in developing diagnostic tests and therapeutic treatments for allergies, asthma, and other immune responses.

The present invention also provides transgenic animals comprising one of the IL4R $\alpha$  genomic polymorphic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the IL4R $\alpha$  isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against IL4R $\alpha$  protein, and for testing the efficacy of therapeutic agents and compounds for allergies, asthma, and other immune responses in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the IL4R $\alpha$  gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the IL4R $\alpha$  gene in a reference

population. In a preferred embodiment, the computer system is capable of producing a display showing IL4R $\alpha$  haplotypes organized according to their evolutionary relationships.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5        Figure 1 illustrates a reference sequence for the IL4R $\alpha$  gene that is the reverse complement of part of Genbank Accession Number AC004525.1; contiguous lines; SEQ ID NO:1), with the underlines indicating the start and stop codons, shading indicating the reference coding sequence, and the bold nucleotides indicating the polymorphic sites and polymorphisms identified by Applicants in a reference population.

10       Figure 2 illustrates a reference sequence for the IL4R $\alpha$  coding sequence (GenBank Accession Number X52425; contiguous lines; SEQ ID NO:2), with the underlines indicating the start and stop codons, and the bold nucleotides indicating the polymorphic sites and polymorphisms identified by Applicants in a reference population.

15       Figure 3 illustrates a reference sequence for the IL4R $\alpha$  protein (GenBank Accession Number CAA36672; contiguous lines; SEQ ID NO:3), with the bold amino acids indicating the amino acid variations caused by the polymorphisms of Fig. 2.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

20       The present invention is based on the discovery of novel variants of the IL4R $\alpha$  gene. As described in more detail below, the inventors herein discovered 38 novel polymorphic sites by characterizing the IL4R $\alpha$  gene found in genomic DNAs isolated from Index Repository IA that contains immortalized cell lines from one chimpanzee and 93 human individuals and Index Repository IB that contains 70 human individuals. These two repositories contain 51 individuals in common.

25       The human individuals in Index Repository IA included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals), African descent (20 individuals) Asian (20 individuals) Hispanic/Latino (17 individuals). To the extent possible, the members of this reference population were organized into population subgroups by the self-identified ethnogeographic origin of their four grandparents as shown in Table 1 below. In addition, Index Repository IA contains three unrelated indigenous American Indians (one from  
30       each of North, Central, and South America), one three-generation Caucasian family (From the CEPH Utah cohort) and one two-generation African-American family.

Table 1. Population Groups in Index Repository IA

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

Index Repository IB contains a reference population of 70 human individuals comprised of 4  
 5 three-generation families (from the CEPH Utah cohort) as well as unrelated African-American, Asian,  
 and Caucasian individuals. A total of 38 individuals in this reference population are unrelated.

Using the IL4R $\alpha$  genotypes identified in the Index Repositories and the methodology described  
 in the Examples below, the inventors herein also determined the haplotypes found on each chromosome  
 for most human members of this repository. The IL4 $\alpha$  genotypes and haplotypes found in the Index  
 10 Repositories include those shown in Tables 4 and 5, respectively. The polymorphism and haplotype data  
 disclosed herein are useful for studying population diversity, anthropological lineage, the significance of  
 diversity and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying  
 associations between the IL4R $\alpha$  genetic variation and a trait such as level of drug response or  
 susceptibility to disease.

15 In the context of this disclosure, the following terms shall be defined as follows unless otherwise  
 indicated:

**Allele** - A particular form of a genetic locus, distinguished from other forms by its particular  
 nucleotide sequence.

**Candidate Gene** - A gene which is hypothesized to be responsible for a disease, condition, or the

response to a treatment, or to be correlated with one of these.

**Gene** - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

5 **Genotype** - An unphased 5' to 3' sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

**Full-genotype** - The unphased 5' to 3' sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

10 **Sub-genotype** - The unphased 5' to 3' sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

**Genotyping** - A process for determining a genotype of an individual.

**Haplotype** - A 5' to 3' sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

15 **Full-haplotype** - The 5' to 3' sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

**Sub-haplotype** - The 5' to 3' sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

**Haplotype pair** - The two haplotypes found for a locus in a single individual.

20 **Haplotyping** - A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

**Haplotype data** - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between  
25 one or more haplotypes and a trait.

**Isoform** - A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

**Isogene** - One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

30 **Isolated** - As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the  
35 present invention.

**Locus** - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

**Naturally- occurring** – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

**Nucleotide pair** – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

**Phased** – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

**Polymorphic site (PS)** – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

**Polymorphic variant** – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

**Polymorphism** – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

**Polymorphism data** – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

**Polymorphism Database** – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

**Polynucleotide** – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

**Population Group** – A group of individuals sharing a common ethnogeographic origin.

**Reference Population** – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

**Single Nucleotide Polymorphism (SNP)** – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

**Subject** – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

**Treatment** – A stimulus administered internally or externally to a subject.

**Unphased** – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.



The inventors herein have discovered 38 novel polymorphic sites, and confirmed the existence of 7 other sites, in the IL4R $\alpha$  gene. The polymorphic sites identified by the inventors are referred to as PS1-45 to designate the order in which they are located in the gene (see Table 3 below), with the novel polymorphic site referred to as PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45 and the previously reported polymorphic sites referred to as PS5, PS24, PS26, PS29, PS31, PS42, and PS43.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the IL4R $\alpha$  gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant IL4R $\alpha$  gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45, and may also comprise one or more additional polymorphisms selected from the group consisting of PS5, PS24, PS26, PS29, PS31, PS42, and PS43. Similarly, the nucleotide sequence of a variant fragment of the IL4R $\alpha$  gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence (or other reported IL4R $\alpha$  sequences) or to portions of the reference sequence (or other reported IL4R $\alpha$  sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of guanine at PS1, thymine at PS2, thymine at PS3, cytosine at PS4, thymine at PS6, adenine at PS7, cytosine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, adenine at PS12, thymine at PS13, thymine at PS14, adenine at PS15, thymine at PS16, adenine at PS17, thymine at PS18, adenine at PS19, cytosine at PS20, cytosine at PS21, thymine at PS22, cytosine at PS23, thymine at PS25, thymine at PS27, cytosine at PS28, thymine at PS30, adenine at PS32, thymine at PS33, guanine at PS34, cytosine at PS35, cytosine at PS36, cytosine at PS37, thymine at PS38, guanine at PS39, guanine at PS40, adenine at PS41, thymine at PS44, and adenine at PS45. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the IL4R $\alpha$  gene which is defined by any one of haplotypes 1-53 shown in Table 5 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the IL4R $\alpha$  gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the polymorphic sites described herein. Any particular variant claimed herein could be

prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

IL4R $\alpha$  isogenes may be isolated using any method that allows separation of the two "copies" of the IL4R $\alpha$  gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and copending U.S. application Serial No. 08/987,966. Another method, which is described in copending U.S. Application Serial No. 08/987,966, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad. Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 17 Nucleic Acids Res. 8392, 1989; Ruaño et al., 19 Nucleic Acids Res. 6877-6882, 1991; Michalatos-Beloin et al., 24 Nucleic Acids Res. 4841-4843, 1996).

The invention also provides IL4R $\alpha$  genome anthologies, which are collections of IL4R $\alpha$  isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. An IL4R $\alpha$  genome anthology may comprise individual IL4R $\alpha$  isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like.

Alternatively, two or more groups of the IL4R $\alpha$  isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred IL4R $\alpha$  genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 5 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded IL4R $\alpha$  protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant IL4R $\alpha$  sequences of the

inv ntion include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the IL4R $\alpha$  gene will produce IL4R $\alpha$  mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of an IL4R $\alpha$  cDNA comprising a nucleotide sequence which is a polymorphic variant of the IL4R $\alpha$  reference coding sequence shown in Figure 2. Thus, the invention also provides IL4R $\alpha$  mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2), or its corresponding RNA sequence, except for having one or more polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 237, adenine at a position corresponding to nucleotide 244, cytosine at a position corresponding to nucleotide 291, thymine at a position corresponding to nucleotide 501, adenine at a position corresponding to nucleotide 554, cytosine at a position corresponding to nucleotide 939, thymine at a position corresponding to nucleotide 1242, thymine at a position corresponding to nucleotide 1293, cytosine at a position corresponding to nucleotide 1299, thymine at a position corresponding to nucleotide 1701, adenine at a position corresponding to nucleotide 1735, thymine at a position corresponding to nucleotide 2023, guanine at a position corresponding to nucleotide 2254 and cytosine at a position corresponding to nucleotide 2397, and may also comprise one or more additional polymorphisms selected from the group consisting of guanine at a position corresponding to 223, cytosine at a position corresponding to nucleotide 1199, cytosine at a position corresponding to 1291, cytosine at a position corresponding to nucleotide 1507 and guanine at a position corresponding to 1737. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized IL4R $\alpha$  cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

Genomic and cDNA fragments of the invention comprise at least one novel polymorphic site identified herein and have a length of at least 10 nucleotides and may range up to the full length of the

gene. Preferably, a fragment according to the present invention is between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the IL4R $\alpha$  gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the IL4R $\alpha$  genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular IL4R $\alpha$  protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the IL4R $\alpha$  isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular IL4R $\alpha$  isogene. Expression of an IL4R $\alpha$  isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3' untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions -10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of IL4R $\alpha$  mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of IL4R $\alpha$  mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2' O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference IL4R $\alpha$  amino acid sequence shown in 3. The location of a variant amino acid in an IL4R $\alpha$  polypeptide or fragment of the invention is identified by aligning its sequence against Fig. 3. An IL4R $\alpha$  protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO: 3 except for having one or more variant amino acids selected from the group consisting of threonine at a position corresponding to amino acid 82, histidine at a position corresponding to amino acid 185, isoleucine at a position corresponding to amino acid 579, serine at a position corresponding to amino acid 675, and alanine at a position corresponding to amino acid 752, and may also comprise one or more additional variant amino acids selected from the group consisting of valine at a position corresponding to amino acid 75, alanine at a position corresponding to amino acid 400, arginine at a position corresponding to amino acid 431, proline at a position corresponding to amino acid 503, and arginine at a position corresponding to amino acid 576. The invention specifically excludes amino acid sequences identical to those previously identified for IL4R $\alpha$ , including SEQ ID NO: 3, and previously described fragments thereof. IL4R $\alpha$  protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO: 3 and having the combination of amino acid variations described in Table 2 below. In preferred embodiments, an IL4R $\alpha$  protein variant of the invention is encoded by an isogene defined by one of the observed haplotypes shown in Table 5.

Table2. Novel Polymorphic Variant of IL4R $\alpha$ 

Polymorphic Variant Number	Amino Acid Position and Id ntities									
	75	82	185	400	431	503	576	579	675	752
1	I	A	R	E	C	S	Q	V	P	A
2	I	A	R	E	C	S	Q	V	S	S
3	I	A	R	E	C	S	Q	V	S	A
4	I	A	R	E	C	S	Q	I	P	S
5	I	A	R	E	C	S	Q	I	P	A
6	I	A	R	E	C	S	Q	I	S	S
7	I	A	R	E	C	S	Q	I	S	A
8	I	A	R	E	C	S	R	V	P	A
9	I	A	R	E	C	S	R	V	S	S
10	I	A	R	E	C	S	R	V	S	A
11	I	A	R	E	C	S	R	I	P	S
12	I	A	R	E	C	S	R	I	P	A
13	I	A	R	E	C	S	R	I	S	S
14	I	A	R	E	C	S	R	I	S	A
15	I	A	R	E	C	P	Q	V	P	A
16	I	A	R	E	C	P	Q	V	S	S
17	I	A	R	E	C	P	Q	V	S	A
18	I	A	R	E	C	P	Q	I	P	S
19	I	A	R	E	C	P	Q	I	P	A
20	I	A	R	E	C	P	Q	I	S	S
21	I	A	R	E	C	P	Q	I	S	A
22	I	A	R	E	C	P	R	V	P	A
23	I	A	R	E	C	P	R	V	S	S
24	I	A	R	E	C	P	R	V	S	A
25	I	A	R	E	C	P	R	I	P	S
26	I	A	R	E	C	P	R	I	P	A
27	I	A	R	E	C	P	R	I	S	S
28	I	A	R	E	C	P	R	I	S	A
29	I	A	R	E	R	S	Q	V	P	A
30	I	A	R	E	R	S	Q	V	S	S
31	I	A	R	E	R	S	Q	V	S	A
32	I	A	R	E	R	S	Q	I	P	S
33	I	A	R	E	R	S	Q	I	P	A
34	I	A	R	E	R	S	Q	I	S	S
35	I	A	R	E	R	S	Q	I	S	A
36	I	A	R	E	R	S	R	V	P	A
37	I	A	R	E	R	S	R	V	S	S
38	I	A	R	E	R	S	R	V	S	A
39	I	A	R	E	R	S	R	I	P	S
40	I	A	R	E	R	S	R	I	P	A
41	I	A	R	E	R	S	R	I	S	S
42	I	A	R	E	R	S	R	I	S	A
43	I	A	R	E	R	P	Q	V	P	A
44	I	A	R	E	R	P	Q	V	S	S
45	I	A	R	E	R	P	Q	V	S	A
46	I	A	R	E	R	P	Q	I	P	S
47	I	A	R	E	R	P	Q	I	P	A
48	I	A	R	E	R	P	Q	I	S	S
49	I	A	R	E	R	P	Q	I	S	A
50	I	A	R	E	R	P	R	V	P	A
51	I	A	R	E	R	P	R	V	S	S
52	I	A	R	E	R	P	R	V	S	A
53	I	A	R	E	R	P	R	I	P	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$ (cont'd)										
Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
54	I	A	R	E	R	P	R	I	P	A
55	I	A	R	E	R	P	R	I	S	S
56	I	A	R	E	R	P	R	I	S	A
57	I	A	R	A	C	S	Q	V	P	A
58	I	A	R	A	C	S	Q	V	S	S
59	I	A	R	A	C	S	Q	V	S	A
60	I	A	R	A	C	S	Q	I	P	S
61	I	A	R	A	C	S	Q	I	P	A
62	I	A	R	A	C	S	Q	I	S	S
63	I	A	R	A	C	S	Q	I	S	A
64	I	A	R	A	C	S	R	V	P	A
65	I	A	R	A	C	S	R	V	S	S
66	I	A	R	A	C	S	R	V	S	A
67	I	A	R	A	C	S	R	I	P	S
68	I	A	R	A	C	S	R	I	P	A
69	I	A	R	A	C	S	R	I	S	S
70	I	A	R	A	C	S	R	I	S	A
71	I	A	R	A	C	P	Q	V	P	A
72	I	A	R	A	C	P	Q	V	S	S
73	I	A	R	A	C	P	Q	V	S	A
74	I	A	R	A	C	P	Q	I	P	S
75	I	A	R	A	C	P	Q	I	P	A
76	I	A	R	A	C	P	Q	I	S	S
77	I	A	R	A	C	P	Q	I	S	A
78	I	A	R	A	C	P	R	V	P	A
79	I	A	R	A	C	P	R	V	S	S
80	I	A	R	A	C	P	R	V	S	A
81	I	A	R	A	C	P	R	I	P	S
82	I	A	R	A	C	P	R	I	P	A
83	I	A	R	A	C	P	R	I	S	S
84	I	A	R	A	C	P	R	I	S	A
85	I	A	R	A	R	S	Q	V	P	A
86	I	A	R	A	R	S	Q	V	S	S
87	I	A	R	A	R	S	Q	V	S	A
88	I	A	R	A	R	S	Q	I	P	S
89	I	A	R	A	R	S	Q	I	P	A
90	I	A	R	A	R	S	Q	I	S	S
91	I	A	R	A	R	S	Q	I	S	A
92	I	A	R	A	R	S	R	V	P	A
93	I	A	R	A	R	S	R	V	S	S
94	I	A	R	A	R	S	R	V	S	A
95	I	A	R	A	R	S	R	I	P	S
96	I	A	R	A	R	S	R	I	P	A
97	I	A	R	A	R	S	R	I	S	S
98	I	A	R	A	R	S	R	I	S	A
99	I	A	R	A	R	P	Q	V	P	A
100	I	A	R	A	R	P	Q	V	S	S
101	I	A	R	A	R	P	Q	V	S	A
102	I	A	R	A	R	P	Q	I	P	S
103	I	A	R	A	R	P	Q	I	P	A
104	I	A	R	A	R	P	Q	I	S	S
105	I	A	R	A	R	P	Q	I	S	A
106	I	A	R	A	R	P	R	V	P	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
107	I	A	R	A	R	P	R	V	S	S
108	I	A	R	A	R	P	R	V	S	A
109	I	A	R	A	R	P	R	I	P	S
110	I	A	R	A	R	P	R	I	P	A
111	I	A	R	A	R	P	R	I	S	S
112	I	A	R	A	R	P	R	I	S	A
113	I	A	H	E	C	S	Q	V	P	S
114	I	A	H	E	C	S	Q	V	P	A
115	I	A	H	E	C	S	Q	V	S	S
116	I	A	H	E	C	S	Q	V	S	A
117	I	A	H	E	C	S	Q	I	P	S
118	I	A	H	E	C	S	Q	I	P	A
119	I	A	H	E	C	S	Q	I	S	S
120	I	A	H	E	C	S	Q	I	S	A
121	I	A	H	E	C	S	R	V	P	S
122	I	A	H	E	C	S	R	V	P	A
123	I	A	H	E	C	S	R	V	S	S
124	I	A	H	E	C	S	R	V	S	A
125	I	A	H	E	C	S	R	I	P	S
126	I	A	H	E	C	S	R	I	P	A
127	I	A	H	E	C	S	R	I	S	S
128	I	A	H	E	C	S	R	I	S	A
129	I	A	H	E	C	P	Q	V	P	S
130	I	A	H	E	C	P	Q	V	P	A
131	I	A	H	E	C	P	Q	V	S	S
132	I	A	H	E	C	P	Q	V	S	A
133	I	A	H	E	C	P	Q	I	P	S
134	I	A	H	E	C	P	Q	I	P	A
135	I	A	H	E	C	P	Q	I	S	S
136	I	A	H	E	C	P	Q	I	S	A
137	I	A	H	E	C	P	R	V	P	S
138	I	A	H	E	C	P	R	V	P	A
139	I	A	H	E	C	P	R	V	S	S
140	I	A	H	E	C	P	R	V	S	A
141	I	A	H	E	C	P	R	I	P	S
142	I	A	H	E	C	P	R	I	P	A
143	I	A	H	E	C	P	R	I	S	S
144	I	A	H	E	C	P	R	I	S	A
145	I	A	H	E	R	S	Q	V	P	S
146	I	A	H	E	R	S	Q	V	P	A
147	I	A	H	E	R	S	Q	V	S	S
148	I	A	H	E	R	S	Q	V	S	A
149	I	A	H	E	R	S	Q	I	P	S
150	I	A	H	E	R	S	Q	I	P	A
151	I	A	H	E	R	S	Q	I	S	S
152	I	A	H	E	R	S	Q	I	S	A
153	I	A	H	E	R	S	R	V	P	S
154	I	A	H	E	R	S	R	V	P	A
155	I	A	H	E	R	S	R	V	S	S
156	I	A	H	E	R	S	R	V	S	A
157	I	A	H	E	R	S	R	I	P	S
158	I	A	H	E	R	S	R	I	P	A
159	I	A	H	E	R	S	R	I	S	S



Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
160	I	A	H	E	R	S	R	I	S	A
161	I	A	H	E	R	P	Q	V	P	S
162	I	A	H	E	R	P	Q	V	P	A
163	I	A	H	E	R	P	Q	V	S	S
164	I	A	H	E	R	P	Q	V	S	A
165	I	A	H	E	R	P	Q	I	P	S
166	I	A	H	E	R	P	Q	I	P	A
167	I	A	H	E	R	P	Q	I	S	S
168	I	A	H	E	R	P	Q	I	S	A
169	I	A	H	E	R	P	R	V	P	S
170	I	A	H	E	R	P	R	V	P	A
171	I	A	H	E	R	P	R	V	S	S
172	I	A	H	E	R	P	R	V	S	A
173	I	A	H	E	R	P	R	I	P	S
174	I	A	H	E	R	P	R	I	P	A
175	I	A	H	E	R	P	R	I	S	S
176	I	A	H	E	R	P	R	I	S	A
177	I	A	H	A	C	S	Q	V	P	S
178	I	A	H	A	C	S	Q	V	P	A
179	I	A	H	A	C	S	Q	V	S	S
180	I	A	H	A	C	S	Q	V	S	A
181	I	A	H	A	C	S	Q	I	P	S
182	I	A	H	A	C	S	Q	I	P	A
183	I	A	H	A	C	S	Q	I	S	S
184	I	A	H	A	C	S	Q	I	S	A
185	I	A	H	A	C	S	R	V	P	S
186	I	A	H	A	C	S	R	V	P	A
187	I	A	H	A	C	S	R	V	S	S
188	I	A	H	A	C	S	R	V	S	A
189	I	A	H	A	C	S	R	I	P	S
190	I	A	H	A	C	S	R	I	P	A
191	I	A	H	A	C	S	R	I	S	S
192	I	A	H	A	C	S	R	I	S	A
193	I	A	H	A	C	P	Q	V	P	S
194	I	A	H	A	C	P	Q	V	P	A
195	I	A	H	A	C	P	Q	V	S	S
196	I	A	H	A	C	P	Q	V	S	A
197	I	A	H	A	C	P	Q	I	P	S
198	I	A	H	A	C	P	Q	I	P	A
199	I	A	H	A	C	P	Q	I	S	S
200	I	A	H	A	C	P	Q	I	S	A
201	I	A	H	A	C	P	R	V	P	S
202	I	A	H	A	C	P	R	V	P	A
203	I	A	H	A	C	P	R	V	S	S
204	I	A	H	A	C	P	R	V	S	A
205	I	A	H	A	C	P	R	I	P	S
206	I	A	H	A	C	P	R	I	P	A
207	I	A	H	A	C	P	R	I	S	S
208	I	A	H	A	C	P	R	I	S	A
209	I	A	H	A	R	S	Q	V	P	S
210	I	A	H	A	R	S	Q	V	P	A
211	I	A	H	A	R	S	Q	V	S	S
212	I	A	H	A	R	S	Q	V	S	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Id ntities									
	75	82	185	400	431	503	576	579	675	752
213	I	A	H	A	R	S	Q	I	P	S
214	I	A	H	A	R	S	Q	I	P	A
215	I	A	H	A	R	S	Q	I	S	S
216	I	A	H	A	R	S	Q	I	S	A
217	I	A	H	A	R	S	R	V	P	S
218	I	A	H	A	R	S	R	V	P	A
219	I	A	H	A	R	S	R	V	S	S
220	I	A	H	A	R	S	R	V	S	A
221	I	A	H	A	R	S	R	I	P	S
222	I	A	H	A	R	S	R	I	P	A
223	I	A	H	A	R	S	R	I	S	S
224	I	A	H	A	R	S	R	I	S	A
225	I	A	H	A	R	P	Q	V	P	S
226	I	A	H	A	R	P	Q	V	P	A
227	I	A	H	A	R	P	Q	V	S	S
228	I	A	H	A	R	P	Q	V	S	A
229	I	A	H	A	R	P	Q	I	P	S
230	I	A	H	A	R	P	Q	I	P	A
231	I	A	H	A	R	P	Q	I	S	S
232	I	A	H	A	R	P	Q	I	S	A
233	I	A	H	A	R	P	R	V	P	S
234	I	A	H	A	R	P	R	V	P	A
235	I	A	H	A	R	P	R	V	S	S
236	I	A	H	A	R	P	R	V	S	A
237	I	A	H	A	R	P	R	I	P	S
238	I	A	H	A	R	P	R	I	P	A
239	I	A	H	A	R	P	R	I	S	S
240	I	A	H	A	R	P	R	I	S	A
241	I	T	R	E	C	S	Q	V	P	S
242	I	T	R	E	C	S	Q	V	P	A
243	I	T	R	E	C	S	Q	V	S	S
244	I	T	R	E	C	S	Q	V	S	A
245	I	T	R	E	C	S	Q	I	P	S
246	I	T	R	E	C	S	Q	I	P	A
247	I	T	R	E	C	S	Q	I	S	S
248	I	T	R	E	C	S	Q	I	S	A
249	I	T	R	E	C	S	R	V	P	S
250	I	T	R	E	C	S	R	V	P	A
251	I	T	R	E	C	S	R	V	S	S
252	I	T	R	E	C	S	R	V	S	A
253	I	T	R	E	C	S	R	I	P	S
254	I	T	R	E	C	S	R	I	P	A
255	I	T	R	E	C	S	R	I	S	S
256	I	T	R	E	C	S	R	I	S	A
257	I	T	R	E	C	P	Q	V	P	S
258	I	T	R	E	C	P	Q	V	P	A
259	I	T	R	E	C	P	Q	V	S	S
260	I	T	R	E	C	P	Q	V	S	A
261	I	T	R	E	C	P	Q	I	P	S
262	I	T	R	E	C	P	Q	I	P	A
263	I	T	R	E	C	P	Q	I	S	S
264	I	T	R	E	C	P	Q	I	S	A
265	I	T	R	E	C	P	R	V	P	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
266	I	T	R	E	C	P	R	V	P	A
267	I	T	R	E	C	P	R	V	S	S
268	I	T	R	E	C	P	R	V	S	A
269	I	T	R	E	C	P	R	I	P	S
270	I	T	R	E	C	P	R	I	P	A
271	I	T	R	E	C	P	R	I	S	S
272	I	T	R	E	C	P	R	I	S	A
273	I	T	R	E	R	S	Q	V	P	S
274	I	T	R	E	R	S	Q	V	P	A
275	I	T	R	E	R	S	Q	V	S	S
276	I	T	R	E	R	S	Q	V	S	A
277	I	T	R	E	R	S	Q	I	P	S
278	I	T	R	E	R	S	Q	I	P	A
279	I	T	R	E	R	S	Q	I	S	S
280	I	T	R	E	R	S	Q	I	S	A
281	I	T	R	E	R	S	R	V	P	S
282	I	T	R	E	R	S	R	V	P	A
283	I	T	R	E	R	S	R	V	S	S
284	I	T	R	E	R	S	R	V	S	A
285	I	T	R	E	R	S	R	I	P	S
286	I	T	R	E	R	S	R	I	P	A
287	I	T	R	E	R	S	R	I	S	S
288	I	T	R	E	R	S	R	I	S	A
289	I	T	R	E	R	P	Q	V	P	S
290	I	T	R	E	R	P	Q	V	P	A
291	I	T	R	E	R	P	Q	V	S	S
292	I	T	R	E	R	P	Q	V	S	A
293	I	T	R	E	R	P	Q	I	P	S
294	I	T	R	E	R	P	Q	I	P	A
295	I	T	R	E	R	P	Q	I	S	S
296	I	T	R	E	R	P	Q	I	S	A
297	I	T	R	E	R	P	R	V	P	S
298	I	T	R	E	R	P	R	V	P	A
299	I	T	R	E	R	P	R	V	S	S
300	I	T	R	E	R	P	R	V	S	A
301	I	T	R	E	R	P	R	I	P	S
302	I	T	R	E	R	P	R	I	P	A
303	I	T	R	E	R	P	R	I	S	S
304	I	T	R	E	R	P	R	I	S	A
305	I	T	R	A	C	S	Q	V	P	S
315	I	T	R	A	C	S	R	V	S	S
316	I	T	R	A	C	S	R	V	S	A
317	I	T	R	A	C	S	R	I	P	S
318	I	T	R	A	C	S	R	I	P	A
319	I	T	R	A	C	S	R	I	S	S
320	I	T	R	A	C	S	R	I	S	A
321	I	T	R	A	C	P	Q	V	P	S
322	I	T	R	A	C	P	Q	V	P	A
323	I	T	R	A	C	P	Q	V	S	S
324	I	T	R	A	C	P	Q	V	S	A
325	I	T	R	A	C	P	Q	I	P	S
326	I	T	R	A	C	P	Q	I	P	A
327	I	T	R	A	C	P	Q	I	S	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
328	I	T	R	A	C	P	Q	I	S	A
306	I	T	R	A	C	S	Q	V	P	A
307	I	T	R	A	C	S	Q	V	S	S
308	I	T	R	A	C	S	Q	V	S	A
309	I	T	R	A	C	S	Q	I	P	S
310	I	T	R	A	C	S	Q	I	P	A
311	I	T	R	A	C	S	Q	I	S	S
312	I	T	R	A	C	S	Q	I	S	A
313	I	T	R	A	C	S	R	V	P	S
314	I	T	R	A	C	S	R	V	P	A
329	I	T	R	A	C	P	R	V	P	S
330	I	T	R	A	C	P	R	V	P	A
331	I	T	R	A	C	P	R	V	S	S
332	I	T	R	A	C	P	R	V	S	A
333	I	T	R	A	C	P	R	I	P	S
334	I	T	R	A	C	P	R	I	P	A
335	I	T	R	A	C	P	R	I	S	S
336	I	T	R	A	C	P	R	I	S	A
337	I	T	R	A	R	S	Q	V	P	S
338	I	T	R	A	R	S	Q	V	P	A
339	I	T	R	A	R	S	Q	V	S	S
340	I	T	R	A	R	S	Q	V	S	A
341	I	T	R	A	R	S	Q	I	P	S
342	I	T	R	A	R	S	Q	I	P	A
343	I	T	R	A	R	S	Q	I	S	S
344	I	T	R	A	R	S	Q	I	S	A
345	I	T	R	A	R	S	R	V	P	S
346	I	T	R	A	R	S	R	V	P	A
347	I	T	R	A	R	S	R	V	S	S
348	I	T	R	A	R	S	R	V	S	A
349	I	T	R	A	R	S	R	I	P	S
350	I	T	R	A	R	S	R	I	P	A
351	I	T	R	A	R	S	R	I	S	S
352	I	T	R	A	R	S	R	I	S	A
353	I	T	R	A	R	P	Q	V	P	S
354	I	T	R	A	R	P	Q	V	P	A
355	I	T	R	A	R	P	Q	V	S	S
356	I	T	R	A	R	P	Q	V	S	A
357	I	T	R	A	R	P	Q	I	P	S
358	I	T	R	A	R	P	Q	I	P	A
359	I	T	R	A	R	P	Q	I	S	S
360	I	T	R	A	R	P	Q	I	S	A
361	I	T	R	A	R	P	R	V	P	S
362	I	T	R	A	R	P	R	V	P	A
363	I	T	R	A	R	P	R	V	S	S
364	I	T	R	A	R	P	R	V	S	A
365	I	T	R	A	R	P	R	I	P	S
366	I	T	R	A	R	P	R	I	P	A
367	I	T	R	A	R	P	R	I	S	S
368	I	T	R	A	R	P	R	I	S	A
369	I	T	H	E	C	S	Q	V	P	S
370	I	T	H	E	C	S	Q	V	P	A
371	I	T	H	E	C	S	Q	V	S	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
372	I	T	H	E	C	S	Q	V	S	A
373	I	T	H	E	C	S	Q	I	P	S
374	I	T	H	E	C	S	Q	I	P	A
375	I	T	H	E	C	S	Q	I	S	S
376	I	T	H	E	C	S	Q	I	S	A
377	I	T	H	E	C	S	R	V	P	S
378	I	T	H	E	C	S	R	V	P	A
379	I	T	H	E	C	S	R	V	S	S
380	I	T	H	E	C	S	R	V	S	A
381	I	T	H	E	C	S	R	I	P	S
382	I	T	H	E	C	S	R	I	P	A
383	I	T	H	E	C	S	R	I	S	S
384	I	T	H	E	C	S	R	I	S	A
385	I	T	H	E	C	P	Q	V	P	S
386	I	T	H	E	C	P	Q	V	P	A
387	I	T	H	E	C	P	Q	V	S	S
388	I	T	H	E	C	P	Q	V	S	A
389	I	T	H	E	C	P	Q	I	P	S
390	I	T	H	E	C	P	Q	I	P	A
391	I	T	H	E	C	P	Q	I	S	S
392	I	T	H	E	C	P	Q	I	S	A
393	I	T	H	E	C	P	R	V	P	S
394	I	T	H	E	C	P	R	V	P	A
395	I	T	H	E	C	P	R	V	S	S
396	I	T	H	E	C	P	R	V	S	A
397	I	T	H	E	C	P	R	I	P	S
398	I	T	H	E	C	P	R	I	P	A
399	I	T	H	E	C	P	R	I	S	S
400	I	T	H	E	C	P	R	I	S	A
401	I	T	H	E	R	S	Q	V	P	S
402	I	T	H	E	R	S	Q	V	P	A
403	I	T	H	E	R	S	Q	V	S	S
404	I	T	H	E	R	S	Q	V	S	A
405	I	T	H	E	R	S	Q	I	P	S
406	I	T	H	E	R	S	Q	I	P	A
407	I	T	H	E	R	S	Q	I	S	S
408	I	T	H	E	R	S	Q	I	S	A
409	I	T	H	E	R	S	R	V	P	S
410	I	T	H	E	R	S	R	V	P	A
411	I	T	H	E	R	S	R	V	S	S
412	I	T	H	E	R	S	R	V	S	A
413	I	T	H	E	R	S	R	I	P	S
414	I	T	H	E	R	S	R	I	P	A
415	I	T	H	E	R	S	R	I	S	S
416	I	T	H	E	R	S	R	I	S	A
417	I	T	H	E	R	P	Q	V	P	S
418	I	T	H	E	R	P	Q	V	P	A
419	I	T	H	E	R	P	Q	V	S	S
420	I	T	H	E	R	P	Q	V	S	A
421	I	T	H	E	R	P	Q	I	P	S
422	I	T	H	E	R	P	Q	I	P	A
423	I	T	H	E	R	P	Q	I	S	S
424	I	T	H	E	R	P	Q	I	S	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
425	I	T	H	E	R	P	R	V	P	S
426	I	T	H	E	R	P	R	V	P	A
427	I	T	H	E	R	P	R	V	S	S
428	I	T	H	E	R	P	R	V	S	A
429	I	T	H	E	R	P	R	I	P	S
430	I	T	H	E	R	P	R	I	P	A
431	I	T	H	E	R	P	R	I	S	S
432	I	T	H	E	R	P	R	I	S	A
433	I	T	H	A	C	S	Q	V	P	S
434	I	T	H	A	C	S	Q	V	P	A
435	I	T	H	A	C	S	Q	V	S	S
436	I	T	H	A	C	S	Q	V	S	A
437	I	T	H	A	C	S	Q	I	P	S
438	I	T	H	A	C	S	Q	I	P	A
439	I	T	H	A	C	S	Q	I	S	S
440	I	T	H	A	C	S	Q	I	S	A
441	I	T	H	A	C	S	R	V	P	S
442	I	T	H	A	C	S	R	V	P	A
443	I	T	H	A	C	S	R	V	S	S
444	I	T	H	A	C	S	R	V	S	A
445	I	T	H	A	C	S	R	I	P	S
446	I	T	H	A	C	S	R	I	P	A
447	I	T	H	A	C	S	R	I	S	S
448	I	T	H	A	C	S	R	I	S	A
449	I	T	H	A	C	P	Q	V	P	S
450	I	T	H	A	C	P	Q	V	P	A
451	I	T	H	A	C	P	Q	V	S	S
452	I	T	H	A	C	P	Q	V	S	A
453	I	T	H	A	C	P	Q	I	P	S
454	I	T	H	A	C	P	Q	I	P	A
455	I	T	H	A	C	P	Q	I	S	S
456	I	T	H	A	C	P	Q	I	S	A
457	I	T	H	A	C	P	R	V	P	S
458	I	T	H	A	C	P	R	V	P	A
459	I	T	H	A	C	P	R	V	S	S
460	I	T	H	A	C	P	R	V	S	A
461	I	T	H	A	C	P	R	I	P	S
462	I	T	H	A	C	P	R	I	P	A
463	I	T	H	A	C	P	R	I	S	S
464	I	T	H	A	C	P	R	I	S	A
465	I	T	H	A	R	S	Q	V	P	S
466	I	T	H	A	R	S	Q	V	P	A
467	I	T	H	A	R	S	Q	V	S	S
468	I	T	H	A	R	S	Q	V	S	A
469	I	T	H	A	R	S	Q	I	P	S
470	I	T	H	A	R	S	Q	I	P	A
471	I	T	H	A	R	S	Q	I	S	S
472	I	T	H	A	R	S	Q	I	S	A
473	I	T	H	A	R	S	R	V	P	S
474	I	T	H	A	R	S	R	V	P	A
475	I	T	H	A	R	S	R	V	S	S
476	I	T	H	A	R	S	R	V	S	A
477	I	T	H	A	R	S	R	I	P	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
478	I	T	H	A	R	S	R	I	P	A
479	I	T	H	A	R	S	R	I	S	S
480	I	T	H	A	R	S	R	I	S	A
481	I	T	H	A	R	P	Q	V	P	S
482	I	T	H	A	R	P	Q	V	P	A
483	I	T	H	A	R	P	Q	V	S	S
484	I	T	H	A	R	P	Q	V	S	A
485	I	T	H	A	R	P	Q	I	P	S
486	I	T	H	A	R	P	Q	I	P	A
487	I	T	H	A	R	P	Q	I	S	S
488	I	T	H	A	R	P	Q	I	S	A
489	I	T	H	A	R	P	R	V	P	S
490	I	T	H	A	R	P	R	V	P	A
491	I	T	H	A	R	P	R	V	S	S
492	I	T	H	A	R	P	R	V	S	A
493	I	T	H	A	R	P	R	I	P	S
494	I	T	H	A	R	P	R	I	P	A
495	I	T	H	A	R	P	R	I	S	S
496	I	T	H	A	R	P	R	I	S	A
497	V	A	R	E	C	S	Q	V	P	A
498	V	A	R	E	C	S	Q	V	S	S
499	V	A	R	E	C	S	Q	V	S	A
500	V	A	R	E	C	S	Q	I	P	S
501	V	A	R	E	C	S	Q	I	P	A
502	V	A	R	E	C	S	Q	I	S	S
503	V	A	R	E	C	S	Q	I	S	A
504	V	A	R	E	C	S	R	V	P	A
505	V	A	R	E	C	S	R	V	S	S
506	V	A	R	E	C	S	R	V	S	A
507	V	A	R	E	C	S	R	I	P	S
508	V	A	R	E	C	S	R	I	P	A
509	V	A	R	E	C	S	R	I	S	S
510	V	A	R	E	C	S	R	I	S	A
511	V	A	R	E	C	P	Q	V	P	A
512	V	A	R	E	C	P	Q	V	S	S
513	V	A	R	E	C	P	Q	V	S	A
514	V	A	R	E	C	P	Q	I	P	S
515	V	A	R	E	C	P	Q	I	P	A
516	V	A	R	E	C	P	Q	I	S	S
517	V	A	R	E	C	P	Q	I	S	A
518	V	A	R	E	C	P	R	V	P	A
519	V	A	R	E	C	P	R	V	S	S
520	V	A	R	E	C	P	R	V	S	A
521	V	A	R	E	C	P	R	I	P	S
522	V	A	R	E	C	P	R	I	P	A
523	V	A	R	E	C	P	R	I	S	S
524	V	A	R	E	C	P	R	I	S	A
525	V	A	R	E	R	S	Q	V	P	A
526	V	A	R	E	R	S	Q	V	S	S
527	V	A	R	E	R	S	Q	V	S	A
528	V	A	R	E	R	S	Q	I	P	S
529	V	A	R	E	R	S	Q	I	P	A
530	V	A	R	E	R	S	Q	I	S	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
531	V	A	R	E	R	S	Q	I	S	A
532	V	A	R	E	R	S	R	V	P	A
533	V	A	R	E	R	S	R	V	S	S
534	V	A	R	E	R	S	R	V	S	A
535	V	A	R	E	R	S	R	I	P	S
536	V	A	R	E	R	S	R	I	P	A
537	V	A	R	E	R	S	R	I	S	S
538	V	A	R	E	R	S	R	I	S	A
539	V	A	R	E	R	P	Q	V	P	A
540	V	A	R	E	R	P	Q	V	S	S
541	V	A	R	E	R	P	Q	V	S	A
542	V	A	R	E	R	P	Q	I	P	S
543	V	A	R	E	R	P	Q	I	P	A
544	V	A	R	E	R	P	Q	I	S	S
545	V	A	R	E	R	P	Q	I	S	A
546	V	A	R	E	R	P	R	V	P	A
547	V	A	R	E	R	P	R	V	S	S
548	V	A	R	E	R	P	R	V	S	A
549	V	A	R	E	R	P	R	I	P	S
550	V	A	R	E	R	P	R	I	P	A
551	V	A	R	E	R	P	R	I	S	S
552	V	A	R	E	R	P	R	I	S	A
553	V	A	R	A	C	S	Q	V	P	A
554	V	A	R	A	C	S	Q	V	S	S
555	V	A	R	A	C	S	Q	V	S	A
556	V	A	R	A	C	S	Q	I	P	S
557	V	A	R	A	C	S	Q	I	P	A
558	V	A	R	A	C	S	Q	I	S	S
559	V	A	R	A	C	S	Q	I	S	A
560	V	A	R	A	C	S	R	V	P	A
561	V	A	R	A	C	S	R	V	S	S
562	V	A	R	A	C	S	R	V	S	A
563	V	A	R	A	C	S	R	I	P	S
564	V	A	R	A	C	S	R	I	P	A
565	V	A	R	A	C	S	R	I	S	S
566	V	A	R	A	C	S	R	I	S	A
567	V	A	R	A	C	P	Q	V	P	A
568	V	A	R	A	C	P	Q	V	S	S
569	V	A	R	A	C	P	Q	V	S	A
570	V	A	R	A	C	P	Q	I	P	S
571	V	A	R	A	C	P	Q	I	P	A
572	V	A	R	A	C	P	Q	I	S	S
573	V	A	R	A	C	P	Q	I	S	A
574	V	A	R	A	C	P	R	V	P	A
575	V	A	R	A	C	P	R	V	S	S
576	V	A	R	A	C	P	R	V	S	A
577	V	A	R	A	C	P	R	I	P	S
578	V	A	R	A	C	P	R	I	P	A
579	V	A	R	A	C	P	R	I	S	S
580	V	A	R	A	C	P	R	I	S	A
581	V	A	R	A	R	S	Q	V	P	A
582	V	A	R	A	R	S	Q	V	S	S
583	V	A	R	A	R	S	Q	V	S	A



Table2. Nov I Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
584	V	A	R	A	R	S	Q	I	P	S
585	V	A	R	A	R	S	Q	I	P	A
586	V	A	R	A	R	S	Q	I	S	S
587	V	A	R	A	R	S	Q	I	S	A
588	V	A	R	A	R	S	R	V	P	A
589	V	A	R	A	R	S	R	V	S	S
590	V	A	R	A	R	S	R	V	S	A
591	V	A	R	A	R	S	R	I	P	S
592	V	A	R	A	R	S	R	I	P	A
593	V	A	R	A	R	S	R	I	S	S
594	V	A	R	A	R	S	R	I	S	A
595	V	A	R	A	R	P	Q	V	P	A
596	V	A	R	A	R	P	Q	V	S	S
597	V	A	R	A	R	P	Q	V	S	A
598	V	A	R	A	R	P	Q	I	P	S
599	V	A	R	A	R	P	Q	I	P	A
600	V	A	R	A	R	P	Q	I	S	S
601	V	A	R	A	R	P	Q	I	S	A
602	V	A	R	A	R	P	R	V	P	A
603	V	A	R	A	R	P	R	V	S	S
604	V	A	R	A	R	P	R	V	S	A
605	V	A	R	A	R	P	R	I	P	S
606	V	A	R	A	R	P	R	I	P	A
607	V	A	R	A	R	P	R	I	S	S
608	V	A	R	A	R	P	R	I	S	A
609	V	A	H	E	C	S	Q	V	P	S
610	V	A	H	E	C	S	Q	V	P	A
611	V	A	H	E	C	S	Q	V	S	S
612	V	A	H	E	C	S	Q	V	S	A
613	V	A	H	E	C	S	Q	I	P	S
614	V	A	H	E	C	S	Q	I	P	A
615	V	A	H	E	C	S	Q	I	S	S
616	V	A	H	E	C	S	Q	I	S	A
617	V	A	H	E	C	S	R	V	P	S
618	V	A	H	E	C	S	R	V	P	A
619	V	A	H	E	C	S	R	V	S	S
620	V	A	H	E	C	S	R	V	S	A
621	V	A	H	E	C	S	R	I	P	S
622	V	A	H	E	C	S	R	I	P	A
623	V	A	H	E	C	S	R	I	S	S
624	V	A	H	E	C	S	R	I	S	A
625	V	A	H	E	C	P	Q	V	P	S
626	V	A	H	E	C	P	Q	V	P	A
627	V	A	H	E	C	P	Q	V	S	S
628	V	A	H	E	C	P	Q	V	S	A
629	V	A	H	E	C	P	Q	I	P	S
630	V	A	H	E	C	P	Q	I	P	A
631	V	A	H	E	C	P	Q	I	S	S
632	V	A	H	E	C	P	Q	I	S	A
633	V	A	H	E	C	P	R	V	P	S
634	V	A	H	E	C	P	R	V	P	A
635	V	A	H	E	C	P	R	V	S	S
636	V	A	H	E	C	P	R	V	S	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
637	V	A	H	E	C	P	R	I	P	S
638	V	A	H	E	C	P	R	I	P	A
639	V	A	H	E	C	P	R	I	S	S
640	V	A	H	E	C	P	R	I	S	A
641	V	A	H	E	R	S	Q	V	P	S
642	V	A	H	E	R	S	Q	V	P	A
643	V	A	H	E	R	S	Q	V	S	S
644	V	A	H	E	R	S	Q	V	S	A
645	V	A	H	E	R	S	Q	I	P	S
646	V	A	H	E	R	S	Q	I	P	A
647	V	A	H	E	R	S	Q	I	S	S
648	V	A	H	E	R	S	Q	I	S	A
649	V	A	H	E	R	S	R	V	P	S
650	V	A	H	E	R	S	R	V	P	A
651	V	A	H	E	R	S	R	V	S	S
652	V	A	H	E	R	S	R	V	S	A
653	V	A	H	E	R	S	R	I	P	S
654	V	A	H	E	R	S	R	I	P	A
655	V	A	H	E	R	S	R	I	S	S
656	V	A	H	E	R	S	R	I	S	A
657	V	A	H	E	R	P	Q	V	P	S
658	V	A	H	E	R	P	Q	V	P	A
659	V	A	H	E	R	P	Q	V	S	S
660	V	A	H	E	R	P	Q	V	S	A
661	V	A	H	E	R	P	Q	I	P	S
662	V	A	H	E	R	P	Q	I	P	A
663	V	A	H	E	R	P	Q	I	S	S
664	V	A	H	E	R	P	Q	I	S	A
665	V	A	H	E	R	P	R	V	P	S
666	V	A	H	E	R	P	R	V	P	A
667	V	A	H	E	R	P	R	V	S	S
668	V	A	H	E	R	P	R	V	S	A
669	V	A	H	E	R	P	R	I	P	S
670	V	A	H	E	R	P	R	I	P	A
671	V	A	H	E	R	P	R	I	S	S
672	V	A	H	E	R	P	R	I	S	A
673	V	A	H	A	C	S	Q	V	P	S
674	V	A	H	A	C	S	Q	V	P	A
675	V	A	H	A	C	S	Q	V	S	S
676	V	A	H	A	C	S	Q	V	S	A
677	V	A	H	A	C	S	Q	I	P	S
678	V	A	H	A	C	S	Q	I	P	A
679	V	A	H	A	C	S	Q	I	S	S
680	V	A	H	A	C	S	Q	I	S	A
681	V	A	H	A	C	S	R	V	P	S
682	V	A	H	A	C	S	R	V	P	A
683	V	A	H	A	C	S	R	V	S	S
684	V	A	H	A	C	S	R	V	S	A
685	V	A	H	A	C	S	R	I	P	S
686	V	A	H	A	C	S	R	I	P	A
687	V	A	H	A	C	S	R	I	S	S
688	V	A	H	A	C	S	R	I	S	A
689	V	A	H	A	C	P	Q	V	P	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
690	V	A	H	A	C	P	Q	V	P	A
691	V	A	H	A	C	P	Q	V	S	S
692	V	A	H	A	C	P	Q	V	S	A
693	V	A	H	A	C	P	Q	I	P	S
694	V	A	H	A	C	P	Q	I	P	A
695	V	A	H	A	C	P	Q	I	S	S
696	V	A	H	A	C	P	Q	I	S	A
697	V	A	H	A	C	P	R	V	P	S
698	V	A	H	A	C	P	R	V	P	A
699	V	A	H	A	C	P	R	V	S	S
700	V	A	H	A	C	P	R	V	S	A
701	V	A	H	A	C	P	R	I	P	S
702	V	A	H	A	C	P	R	I	P	A
703	V	A	H	A	C	P	R	I	S	S
704	V	A	H	A	C	P	R	I	S	A
705	V	A	H	A	R	S	Q	V	P	S
706	V	A	H	A	R	S	Q	V	P	A
707	V	A	H	A	R	S	Q	V	S	S
708	V	A	H	A	R	S	Q	V	S	A
709	V	A	H	A	R	S	Q	I	P	S
710	V	A	H	A	R	S	Q	I	P	A
711	V	A	H	A	R	S	Q	I	S	S
712	V	A	H	A	R	S	Q	I	S	A
713	V	A	H	A	R	S	R	V	P	S
714	V	A	H	A	R	S	R	V	P	A
715	V	A	H	A	R	S	R	V	S	S
716	V	A	H	A	R	S	R	V	S	A
717	V	A	H	A	R	S	R	I	P	S
718	V	A	H	A	R	S	R	I	P	A
719	V	A	H	A	R	S	R	I	S	S
720	V	A	H	A	R	S	R	I	S	A
721	V	A	H	A	R	P	Q	V	P	S
722	V	A	H	A	R	P	Q	V	P	A
723	V	A	H	A	R	P	Q	V	S	S
724	V	A	H	A	R	P	Q	V	S	A
725	V	A	H	A	R	P	Q	I	P	S
726	V	A	H	A	R	P	Q	I	P	A
727	V	A	H	A	R	P	Q	I	S	S
728	V	A	H	A	R	P	Q	I	S	A
729	V	A	H	A	R	P	R	V	P	S
730	V	A	H	A	R	P	R	V	P	A
731	V	A	H	A	R	P	R	V	S	S
732	V	A	H	A	R	P	R	V	S	A
733	V	A	H	A	R	P	R	I	P	S
734	V	A	H	A	R	P	R	I	P	A
735	V	A	H	A	R	P	R	I	S	S
736	V	A	H	A	R	P	R	I	S	A
737	V	T	R	E	C	S	Q	V	P	S
738	V	T	R	E	C	S	Q	V	P	A
739	V	T	R	E	C	S	Q	V	S	S
740	V	T	R	E	C	S	Q	V	S	A
741	V	T	R	E	C	S	Q	I	P	S
742	V	T	R	E	C	S	Q	I	P	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
743	V	T	R	E	C	S	Q	I	S	S
744	V	T	R	E	C	S	Q	I	S	A
745	V	T	R	E	C	S	R	V	P	S
746	V	T	R	E	C	S	R	V	P	A
747	V	T	R	E	C	S	R	V	S	S
748	V	T	R	E	C	S	R	V	S	A
749	V	T	R	E	C	S	R	I	P	S
750	V	T	R	E	C	S	R	I	P	A
751	V	T	R	E	C	S	R	I	S	S
752	V	T	R	E	C	S	R	I	S	A
753	V	T	R	E	C	P	Q	V	P	S
754	V	T	R	E	C	P	Q	V	P	A
755	V	T	R	E	C	P	Q	V	S	S
756	V	T	R	E	C	P	Q	V	S	A
757	V	T	R	E	C	P	Q	I	P	S
758	V	T	R	E	C	P	Q	I	P	A
759	V	T	R	E	C	P	Q	I	S	S
760	V	T	R	E	C	P	Q	I	S	A
761	V	T	R	E	C	P	R	V	P	S
762	V	T	R	E	C	P	R	V	P	A
763	V	T	R	E	C	P	R	V	S	S
764	V	T	R	E	C	P	R	V	S	A
765	V	T	R	E	C	P	R	I	P	S
766	V	T	R	E	C	P	R	I	P	A
767	V	T	R	E	C	P	R	I	S	S
768	V	T	R	E	C	P	R	I	S	A
769	V	T	R	E	R	S	Q	V	P	S
770	V	T	R	E	R	S	Q	V	P	A
771	V	T	R	E	R	S	Q	V	S	S
772	V	T	R	E	R	S	Q	V	S	A
773	V	T	R	E	R	S	Q	I	P	S
774	V	T	R	E	R	S	Q	I	P	A
775	V	T	R	E	R	S	Q	I	S	S
776	V	T	R	E	R	S	Q	I	S	A
777	V	T	R	E	R	S	R	V	P	S
778	V	T	R	E	R	S	R	V	P	A
779	V	T	R	E	R	S	R	V	S	S
780	V	T	R	E	R	S	R	V	S	A
781	V	T	R	E	R	S	R	I	P	S
782	V	T	R	E	R	S	R	I	P	A
783	V	T	R	E	R	S	R	I	S	S
784	V	T	R	E	R	S	R	I	S	A
785	V	T	R	E	R	P	Q	V	P	S
786	V	T	R	E	R	P	Q	V	P	A
787	V	T	R	E	R	P	Q	V	S	S
788	V	T	R	E	R	P	Q	V	S	A
789	V	T	R	E	R	P	Q	I	P	S
790	V	T	R	E	R	P	Q	I	P	A
791	V	T	R	E	R	P	Q	I	S	S
792	V	T	R	E	R	P	Q	I	S	A
793	V	T	R	E	R	P	R	V	P	S
794	V	T	R	E	R	P	R	V	P	A
795	V	T	R	E	R	P	R	V	S	S

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
796	V	T	R	E	R	P	R	V	S	A
797	V	T	R	E	R	P	R	I	P	S
798	V	T	R	E	R	P	R	I	P	A
799	V	T	R	E	R	P	R	I	S	S
800	V	T	R	E	R	P	R	I	S	A
801	V	T	R	A	C	S	Q	V	P	S
802	V	T	R	A	C	S	Q	V	P	A
803	V	T	R	A	C	S	Q	V	S	S
804	V	T	R	A	C	S	Q	V	S	A
805	V	T	R	A	C	S	Q	I	P	S
806	V	T	R	A	C	S	Q	I	P	A
807	V	T	R	A	C	S	Q	I	S	S
808	V	T	R	A	C	S	Q	I	S	A
809	V	T	R	A	C	S	R	V	P	S
810	V	T	R	A	C	S	R	V	P	A
811	V	T	R	A	C	S	R	V	S	S
812	V	T	R	A	C	S	R	V	S	A
813	V	T	R	A	C	S	R	I	P	S
814	V	T	R	A	C	S	R	I	P	A
815	V	T	R	A	C	S	R	I	S	S
816	V	T	R	A	C	S	R	I	S	A
817	V	T	R	A	C	P	Q	V	P	S
818	V	T	R	A	C	P	Q	V	P	A
819	V	T	R	A	C	P	Q	V	S	S
820	V	T	R	A	C	P	Q	V	S	A
821	V	T	R	A	C	P	Q	I	P	S
822	V	T	R	A	C	P	Q	I	P	A
823	V	T	R	A	C	P	Q	I	S	S
824	V	T	R	A	C	P	Q	I	S	A
825	V	T	R	A	C	P	R	V	P	S
826	V	T	R	A	C	P	R	V	P	A
827	V	T	R	A	C	P	R	V	S	S
828	V	T	R	A	C	P	R	V	S	A
829	V	T	R	A	C	P	R	I	P	S
830	V	T	R	A	C	P	R	I	P	A
831	V	T	R	A	C	P	R	I	S	S
832	V	T	R	A	C	P	R	I	S	A
833	V	T	R	A	R	S	Q	V	P	S
834	V	T	R	A	R	S	Q	V	P	A
835	V	T	R	A	R	S	Q	V	S	S
836	V	T	R	A	R	S	Q	V	S	A
837	V	T	R	A	R	S	Q	I	P	S
838	V	T	R	A	R	S	Q	I	P	A
839	V	T	R	A	R	S	Q	I	S	S
840	V	T	R	A	R	S	Q	I	S	A
841	V	T	R	A	R	S	R	V	P	S
842	V	T	R	A	R	S	R	V	P	A
843	V	T	R	A	R	S	R	V	S	S
844	V	T	R	A	R	S	R	V	S	A
845	V	T	R	A	R	S	R	I	P	S
846	V	T	R	A	R	S	R	I	P	A
847	V	T	R	A	R	S	R	I	S	S
848	V	T	R	A	R	S	R	I	S	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
849	V	T	R	A	R	P	Q	V	P	S
850	V	T	R	A	R	P	Q	V	P	A
851	V	T	R	A	R	P	Q	V	S	S
852	V	T	R	A	R	P	Q	V	S	A
853	V	T	R	A	R	P	Q	I	P	S
854	V	T	R	A	R	P	Q	I	P	A
855	V	T	R	A	R	P	Q	I	S	S
856	V	T	R	A	R	P	Q	I	S	A
857	V	T	R	A	R	P	R	V	P	S
858	V	T	R	A	R	P	R	V	P	A
859	V	T	R	A	R	P	R	V	S	S
860	V	T	R	A	R	P	R	V	S	A
861	V	T	R	A	R	P	R	I	P	S
862	V	T	R	A	R	P	R	I	P	A
863	V	T	R	A	R	P	R	I	S	S
864	V	T	R	A	R	P	R	I	S	A
865	V	T	H	E	C	S	Q	V	P	S
866	V	T	H	E	C	S	Q	V	P	A
867	V	T	H	E	C	S	Q	V	S	S
868	V	T	H	E	C	S	Q	V	S	A
869	V	T	H	E	C	S	Q	I	P	S
870	V	T	H	E	C	S	Q	I	P	A
871	V	T	H	E	C	S	Q	I	S	S
872	V	T	H	E	C	S	Q	I	S	A
873	V	T	H	E	C	S	R	V	P	S
874	V	T	H	E	C	S	R	V	P	A
875	V	T	H	E	C	S	R	V	S	S
876	V	T	H	E	C	S	R	V	S	A
877	V	T	H	E	C	S	R	I	P	S
878	V	T	H	E	C	S	R	I	P	A
879	V	T	H	E	C	S	R	I	S	S
880	V	T	H	E	C	S	R	I	S	A
881	V	T	H	E	C	P	Q	V	P	S
882	V	T	H	E	C	P	Q	V	P	A
883	V	T	H	E	C	P	Q	V	S	S
884	V	T	H	E	C	P	Q	V	S	A
885	V	T	H	E	C	P	Q	I	P	S
886	V	T	H	E	C	P	Q	I	P	A
887	V	T	H	E	C	P	Q	I	S	S
888	V	T	H	E	C	P	Q	I	S	A
889	V	T	H	E	C	P	R	V	P	S
890	V	T	H	E	C	P	R	V	P	A
891	V	T	H	E	C	P	R	V	S	S
892	V	T	H	E	C	P	R	V	S	A
893	V	T	H	E	C	P	R	I	P	S
894	V	T	H	E	C	P	R	I	P	A
895	V	T	H	E	C	P	R	I	S	S
896	V	T	H	E	C	P	R	I	S	A
897	V	T	H	E	R	S	Q	V	P	S
898	V	T	H	E	R	S	Q	V	P	A
899	V	T	H	E	R	S	Q	V	S	S
900	V	T	H	E	R	S	Q	V	S	A
901	V	T	H	E	R	S	Q	I	P	S

Tabl 2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
902	V	T	H	E	R	S	Q	I	P	A
903	V	T	H	E	R	S	Q	I	S	S
904	V	T	H	E	R	S	Q	I	S	A
905	V	T	H	E	R	S	R	V	P	S
906	V	T	H	E	R	S	R	V	P	A
907	V	T	H	E	R	S	R	V	S	S
908	V	T	H	E	R	S	R	V	S	A
909	V	T	H	E	R	S	R	I	P	S
910	V	T	H	E	R	S	R	I	P	A
911	V	T	H	E	R	S	R	I	S	S
912	V	T	H	E	R	S	R	I	S	A
913	V	T	H	E	R	P	Q	V	P	S
914	V	T	H	E	R	P	Q	V	P	A
915	V	T	H	E	R	P	Q	V	S	S
916	V	T	H	E	R	P	Q	V	S	A
917	V	T	H	E	R	P	Q	I	P	S
918	V	T	H	E	R	P	Q	I	P	A
919	V	T	H	E	R	P	Q	I	S	S
920	V	T	H	E	R	P	Q	I	S	A
921	V	T	H	E	R	P	R	V	P	S
922	V	T	H	E	R	P	R	V	P	A
923	V	T	H	E	R	P	R	V	S	S
924	V	T	H	E	R	P	R	V	S	A
925	V	T	H	E	R	P	R	I	P	S
926	V	T	H	E	R	P	R	I	P	A
927	V	T	H	E	R	P	R	I	S	S
928	V	T	H	E	R	P	R	I	S	A
929	V	T	H	A	C	S	Q	V	P	S
930	V	T	H	A	C	S	Q	V	P	A
931	V	T	H	A	C	S	Q	V	S	S
932	V	T	H	A	C	S	Q	V	S	A
933	V	T	H	A	C	S	Q	I	P	S
934	V	T	H	A	C	S	Q	I	P	A
935	V	T	H	A	C	S	Q	I	S	S
936	V	T	H	A	C	S	Q	I	S	A
937	V	T	H	A	C	S	R	V	P	S
938	V	T	H	A	C	S	R	V	P	A
939	V	T	H	A	C	S	R	V	S	S
940	V	T	H	A	C	S	R	V	S	A
941	V	T	H	A	C	S	R	I	P	S
942	V	T	H	A	C	S	R	I	P	A
943	V	T	H	A	C	S	R	I	S	S
944	V	T	H	A	C	S	R	I	S	A
945	V	T	H	A	C	P	Q	V	P	S
946	V	T	H	A	C	P	Q	V	P	A
947	V	T	H	A	C	P	Q	V	S	S
948	V	T	H	A	C	P	Q	V	S	A
949	V	T	H	A	C	P	Q	I	P	S
950	V	T	H	A	C	P	Q	I	P	A
951	V	T	H	A	C	P	Q	I	S	S
952	V	T	H	A	C	P	Q	I	S	A
953	V	T	H	A	C	P	R	V	P	S
954	V	T	H	A	C	P	R	V	P	A

Table2. Novel Polymorphic Variant of IL4R $\alpha$  (cont'd)

Polymorphic Variant Number	Amino Acid Position and Identities									
	75	82	185	400	431	503	576	579	675	752
955	V	T	H	A	C	P	R	V	S	S
956	V	T	H	A	C	P	R	V	S	A
957	V	T	H	A	C	P	R	I	P	S
958	V	T	H	A	C	P	R	I	P	A
959	V	T	H	A	C	P	R	I	S	S
960	V	T	H	A	C	P	R	I	S	A
961	V	T	H	A	R	S	Q	V	P	S
962	V	T	H	A	R	S	Q	V	P	A
963	V	T	H	A	R	S	Q	V	S	S
964	V	T	H	A	R	S	Q	V	S	A
965	V	T	H	A	R	S	Q	I	P	S
966	V	T	H	A	R	S	Q	I	P	A
967	V	T	H	A	R	S	Q	I	S	S
968	V	T	H	A	R	S	Q	I	S	A
969	V	T	H	A	R	S	R	V	P	S
970	V	T	H	A	R	S	R	V	P	A
971	V	T	H	A	R	S	R	V	S	S
972	V	T	H	A	R	S	R	V	S	A
973	V	T	H	A	R	S	R	I	P	S
974	V	T	H	A	R	S	R	I	P	A
975	V	T	H	A	R	S	R	I	S	S
976	V	T	H	A	R	S	R	I	S	A
977	V	T	H	A	R	P	Q	V	P	S
978	V	T	H	A	R	P	Q	V	P	A
979	V	T	H	A	R	P	Q	V	S	S
980	V	T	H	A	R	P	Q	V	S	A
981	V	T	H	A	R	P	Q	I	P	S
982	V	T	H	A	R	P	Q	I	P	A
983	V	T	H	A	R	P	Q	I	S	S
984	V	T	H	A	R	P	Q	I	S	A
985	V	T	H	A	R	P	R	V	P	S
986	V	T	H	A	R	P	R	V	P	A
987	V	T	H	A	R	P	R	V	S	S
988	V	T	H	A	R	P	R	V	S	A
989	V	T	H	A	R	P	R	I	P	S
990	V	T	H	A	R	P	R	I	P	A
991	V	T	H	A	R	P	R	I	S	S
992	V	T	H	A	R	P	R	I	S	A

The invention also includes IL4R $\alpha$  peptide variants, which are any fragments of an IL4R $\alpha$  protein variant that contains one or more of the amino acid variations shown in Table 2. An IL4R $\alpha$  peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such IL4R $\alpha$  peptide variants may be useful as antigens to generate antibodies specific for one of the above IL4R $\alpha$  isoforms. In addition, the IL4R $\alpha$  peptide variants may be useful in drug screening assays.

An IL4R $\alpha$  variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant IL4R $\alpha$  genomic and cDNA sequences as described above. Alternatively, the IL4R $\alpha$  protein variant may be isolated from a biological sample of an individual having an IL4R $\alpha$



isogene which encodes the variant protein. Where the sample contains two different IL4R $\alpha$  isoforms (i.e., the individual has different IL4R $\alpha$  isogenes), a particular IL4R $\alpha$  isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular IL4R $\alpha$  isoform but does not bind to the other IL4R $\alpha$  isoform.

5       The expressed or isolated IL4R $\alpha$  protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the IL4R $\alpha$  protein as discussed further below. IL4R $\alpha$  variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity  
10       and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant IL4R $\alpha$  gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric IL4R $\alpha$  protein. The non-IL4R $\alpha$  portion of the chimeric  
15       protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the IL4R $\alpha$  and non-IL4R $\alpha$  portions so that the IL4R $\alpha$  protein may be cleaved and purified away from the non-IL4R $\alpha$  portion.

An additional embodiment of the invention relates to using a novel IL4R $\alpha$  protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that  
20       bind specifically to all known IL4R $\alpha$  protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The IL4R $\alpha$  protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to an IL4R $\alpha$  variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of  
25       test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the IL4R $\alpha$  protein(s) of interest and then washed. Bound IL4R $\alpha$  protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel IL4R $\alpha$  protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the IL4R $\alpha$  protein.

30       In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel IL4R $\alpha$  variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The IL4R $\alpha$  protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the IL4R $\alpha$  protein variant is of insufficient size to be antigenic, it may be  
35       conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human.

bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel IL4R $\alpha$  protein isoforms described herein is administered to an individual to neutralize activity of the IL4R $\alpha$  isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel IL4R $\alpha$  protein isoform described herein may be used to immunoprecipitate the IL4R $\alpha$  protein variant from solution as well as react with IL4R $\alpha$  protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect IL4R $\alpha$  protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel IL4R $\alpha$  protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the IL4R $\alpha$  protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The

antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in *E. coli* is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci. 86:10029).

Effect(s) of the polymorphisms identified herein on expression of IL4R $\alpha$  may be investigated by preparing recombinant cells and/or organisms, preferably recombinant animals, containing a polymorphic variant of the IL4R $\alpha$  gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into IL4R $\alpha$  protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired IL4R $\alpha$  isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the IL4R $\alpha$  isogene is introduced into a cell in such a way that it recombines with the endogenous IL4R $\alpha$  gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired IL4R $\alpha$  gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the IL4R $\alpha$  isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the IL4R $\alpha$  isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant organisms, i.e., transgenic animals, expressing a variant IL4R $\alpha$  gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the IL4R $\alpha$  isogenes may be introduced include,

but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human IL4R $\alpha$  isogene and producing human IL4R $\alpha$  protein can be used as biological models for studying diseases related to abnormal IL4R $\alpha$  expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel IL4R $\alpha$  isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel IL4R $\alpha$  isogenes; an antisense oligonucleotide directed against one of the novel IL4R $\alpha$  isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel IL4R $\alpha$  isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel IL4R $\alpha$  isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Information on the identity of genotypes and haplotypes for the IL4R $\alpha$  gene of any particular

individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is expected to be useful for a variety of basic research and clinical applications. Thus, the invention also provides compositions and methods for detecting the novel IL4R $\alpha$  polymorphisms identified herein.

5 The compositions comprise at least one IL4R $\alpha$  genotyping oligonucleotide. In one embodiment, an IL4R $\alpha$  genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the  
10 oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R.  
15 in *Molecular Biology and Biotechnology*, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens,  
20 antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of an IL4R $\alpha$  polynucleotide, i.e., an IL4R $\alpha$  isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target  
25 region or a non-IL4R $\alpha$  polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the IL4R $\alpha$  gene using the polymorphism information provided herein in conjunction with the known sequence information for the IL4R $\alpha$  gene and routine techniques.

30 A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions.

35 Conventional hybridization conditions are described, for example, by Sambrook J. et al., in *Molecular Cloning*, A Laboratory Manual, 2<sup>nd</sup> Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989)

and by Haymes, B.D. et al. in *Nucleic Acid Hybridization, A Practical Approach*, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5' end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in *PCR Protocols, A Guide to Methods and Applications*, Academic Press, 1990 and Ruano et al., 87 *Proc. Natl. Acad. Sci. USA* 6296-6300, 1990. Typically, an allele-specific oligonucleotide will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotide probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7<sup>th</sup> or 8<sup>th</sup> position in a 15 mer, the 8<sup>th</sup> or 9<sup>th</sup> position in a 16mer, the 10<sup>th</sup> or 11<sup>th</sup> position in a 20 mer). A preferred ASO probe for detecting IL4R $\alpha$  gene polymorphisms comprises a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

TTGCACCACTGCACT	(SEQ ID NO:4) and its complement,
TTGCACCGCTGCACT	(SEQ ID NO:5) and its complement,
TTTTGTGCTATTCCC	(SEQ ID NO:6) and its complement,
TTTTGTGTTATTCCC	(SEQ ID NO:7) and its complement,
CTGGGCCGCTCAGGC	(SEQ ID NO:8) and its complement,
CTGGGCCTCTCAGGC	(SEQ ID NO:9) and its complement,
TAAGCCTGCGCTGGA	(SEQ ID NO:10) and its complement,
TAAGCCTCCGCTGGA	(SEQ ID NO:11) and its complement,
AGAACAACGGAGGCG	(SEQ ID NO:12) and its complement,
AGAACAATGGAGGCG	(SEQ ID NO:13) and its complement,
CGGAGGCGCGGGGTG	(SEQ ID NO:14) and its complement,
CGGAGGCACGGGGTG	(SEQ ID NO:15) and its complement,

GTGCGGATAACTATA (SEQ ID NO:16) and its complement,  
 GTGCGGACAACATA (SEQ ID NO:17) and its complement,  
 5 CGGAGTGCGGCAGGG (SEQ ID NO:18) and its complement,  
 CGGAGTGTGGCAGGG (SEQ ID NO:19) and its complement,  
 GCCTGGGCTGAGGGT (SEQ ID NO:20) and its complement,  
 GCCTGGGTTGAGGGT (SEQ ID NO:21) and its complement,  
 10 TGGGGTGCGGCAGGGG (SEQ ID NO:22) and its complement,  
 TGGGGTGAGCAGGGG (SEQ ID NO:23) and its complement,  
 TTCTCCCGCAGTGAA (SEQ ID NO:24) and its complement,  
 15 TTCTCCACAGTGAA (SEQ ID NO:25) and its complement,  
 GTGAAAACGACCCGG (SEQ ID NO:26) and its complement,  
 GTGAAAATGACCCGG (SEQ ID NO:27) and its complement,  
 GGCAAGCCCTGGGGC (SEQ ID NO:28) and its complement,  
 20 GGCAAGCTCTGGGGC (SEQ ID NO:29) and its complement,  
 GCCCTGGGGCTGGAT (SEQ ID NO:30) and its complement,  
 GCCCTGGAGCTGGAT (SEQ ID NO:31) and its complement,  
 25 ATAGCAAATCCCAGG (SEQ ID NO:32) and its complement,  
 ATAGCAATTCCCAGG (SEQ ID NO:33) and its complement,  
 GCTCTGCCCTAGGCA (SEQ ID NO:34) and its complement,  
 30 GCTCTGCACTAGGCA (SEQ ID NO:35) and its complement,  
 CCCCCACCCCTCACA (SEQ ID NO:36) and its complement,  
 CCCCCACTCCTCACA (SEQ ID NO:37) and its complement,  
 TCCCTCCGCATCGCA (SEQ ID NO:38) and its complement,  
 35 TCCCTCCACATCGCA (SEQ ID NO:39) and its complement,  
 CACCTGCTGTGGTGT (SEQ ID NO:40) and its complement,  
 CACCTGCCGTGGTGT (SEQ ID NO:41) and its complement,  
 40 ATGTCTGAAGTAGAC (SEQ ID NO:42) and its complement,  
 ATGTCTGCAGTAGAC (SEQ ID NO:43) and its complement,  
 TGACCAACCTTTGCT (SEQ ID NO:44) and its complement,  
 45 TGACCAACTTTTGCT (SEQ ID NO:45) and its complement,  
 CCTGTTTTCTGGAGC (SEQ ID NO:46) and its complement,  
 CCTGTTTCCTGGAGC (SEQ ID NO:47) and its complement,  
 TGGACCTGCTCGGAG (SEQ ID NO:48) and its complement,  
 50 TGGACCTTCTCGGAG (SEQ ID NO:49) and its complement,  
 AGTCATGCCTTCTTC (SEQ ID NO:50) and its complement,  
 AGTCATGTCTTCTTC (SEQ ID NO:51) and its complement,  
 55 GCCTTCTTCCACCTT (SEQ ID NO:52) and its complement,  
 GCCTTCTCCACCTT (SEQ ID NO:53) and its complement,

CAGCCCCCGTCTCGG (SEQ ID NO:54) and its complement,  
 CAGCCCCGTGTCTCGG (SEQ ID NO:55) and its complement,  
 5 GGAGTTTGTACATGC (SEQ ID NO:56) and its complement,  
 GGAGTTTATACATGC (SEQ ID NO:57) and its complement,  
 CAGCTCCCCAGAGCA (SEQ ID NO:58) and its complement,  
 CAGCTCCTCAGAGCA (SEQ ID NO:59) and its complement,  
 10 AGACAGGTCCTCGCC (SEQ ID NO:60) and its complement,  
 AGACAGGGCCTCGCC (SEQ ID NO:61) and its complement,  
 CTGCCCCCTGGCAATG (SEQ ID NO:62) and its complement,  
 CTGCCCCCGGCAATG (SEQ ID NO:63) and its complement,  
 15 AGGTGCATGTCCTCT (SEQ ID NO:64) and its complement,  
 AGGTGCACGTCCTCT (SEQ ID NO:65) and its complement,  
 GTGCATGTCCTCTTG (SEQ ID NO:66) and its complement,  
 20 GTGCATGCCCTCTTG (SEQ ID NO:67) and its complement,  
 GGCTTATCCATGCCT (SEQ ID NO:68) and its complement,  
 GGCTTATTCATGCCT (SEQ ID NO:69) and its complement,  
 25 AGCCAGGCTGGCAGA (SEQ ID NO:70) and its complement,  
 AGCCAGGGTGGCAGA (SEQ ID NO:71) and its complement,  
 GGCCCACATGGAGGC (SEQ ID NO:72) and its complement,  
 GGCCCACGTGGAGGC (SEQ ID NO:73) and its complement,  
 30 TAACACAGCCATCAA (SEQ ID NO:74) and its complement,  
 TAACACAACCATCAA (SEQ ID NO:75) and its complement,  
 TAATGCTCGTCTGTG (SEQ ID NO:76) and its complement,  
 35 TAATGCTTGTCTGTG (SEQ ID NO:77) and its complement,  
 ACTTGCCGTCTGGGT (SEQ ID NO:78) and its complement, and  
 ACTTGCCATCTGGGT (SEQ ID NO:79) and its complement.  
 40

An allele-specific oligonucleotide primer of the invention has a 3' terminal nucleotide, or preferably a 3' penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. Allele-specific oligonucleotide primers hybridizing to either the coding or noncoding strand

45 are contemplated by the invention. A preferred ASO primer for detecting IL4R $\alpha$  gene polymorphisms comprises a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

CTGAGATTGCACCAC (SEQ ID NO:80); GGCTGGAGTGCAGTG (SEQ ID NO:81);  
 CTGAGATTGCACCGC (SEQ ID NO:82); GGCTGGAGTGCAGCG (SEQ ID NO:83);

50 CTGTGCTTTTGTGCT (SEQ ID NO:84); ACCAAGGGGAATAGC (SEQ ID NO:85);  
 CTGTGCTTTTGTGTT (SEQ ID NO:86); ACCAAGGGGAATAAC (SEQ ID NO:87);



GAGTTCCTGGGCCGC (SEQ ID NO:88); GGAGCAGCCTGAGCG (SEQ ID NO:89);  
GAGTTCCTGGGCCTC (SEQ ID NO:90); GGAGCAGCCTGAGAG (SEQ ID NO:91);

5 TCCGAGTAAGCCTGC (SEQ ID NO:92); TCCAGCTCCAGCGCA (SEQ ID NO:93);  
TCCGAGTAAGCCTCC (SEQ ID NO:94); TCCAGCTCCAGCGGA (SEQ ID NO:95);

TCCCTGAGAACAACG (SEQ ID NO:96); ACCCCGCGCCTCCGT (SEQ ID NO:97);  
TCCCTGAGAACAATG (SEQ ID NO:98); ACCCCGCGCCTCCAT (SEQ ID NO:99);

10 GAACAACGGAGGCGC (SEQ ID NO:100); CACACGCACCCCGCG (SEQ ID NO:101);  
GAACAACGGAGGCAC (SEQ ID NO:102); CACACGCACCCCGTG (SEQ ID NO:103);

TGGTCAGTGCGGATA (SEQ ID NO:104); CCAGTGTATAGTTAT (SEQ ID NO:105);  
15 TGGTCAGTGCGGACA (SEQ ID NO:106); CCAGTGTATAGTTGT (SEQ ID NO:107);

GCAGGGCGGAGTGCG (SEQ ID NO:108); AGCCACCCCTGCCGC (SEQ ID NO:109);  
GCAGGGCGGAGTGTG (SEQ ID NO:110); AGCCACCCCTGCCAC (SEQ ID NO:111);

20 ACAGCTGCCTGGGCT (SEQ ID NO:112); CACCCACCCCTCAGC (SEQ ID NO:113);  
ACAGCTGCCTGGGTT (SEQ ID NO:114); CACCCACCCCTCAAC (SEQ ID NO:115);

TGAGGGTGGGGTGGG (SEQ ID NO:116); CCTCCTCCCCTGCCC (SEQ ID NO:117);  
TGAGGGTGGGGTGAG (SEQ ID NO:118); CCTCCTCCCCTGCTC (SEQ ID NO:119);

25 GGCCGCTTCTCCCGC (SEQ ID NO:120); CTGGGTTTCACTGCG (SEQ ID NO:121);  
GGCCGCTTCTCCAC (SEQ ID NO:122); CTGGGTTTCACTGTG (SEQ ID NO:123);

30 TTTGGAGTGAAACG (SEQ ID NO:124); CATCTGCCGGGTCGT (SEQ ID NO:125);  
TTTGGAGTGAAATG (SEQ ID NO:126); CATCTGCCGGGTCAT (SEQ ID NO:127);

CTGGGAGGCAAGCCC (SEQ ID NO:128); TATCCAGCCCCAGGG (SEQ ID NO:129);  
CTGGGAGGCAAGCTC (SEQ ID NO:130); TATCCAGCCCCAGAG (SEQ ID NO:131);

35 AGGCAAGCCCTGGGG (SEQ ID NO:132); TTTGCTATCCAGCCC (SEQ ID NO:133);  
AGGCAAGCCCTGGAG (SEQ ID NO:134); TTTGCTATCCAGCTC (SEQ ID NO:135);

GGCTGGATAGCAAAT (SEQ ID NO:136); CTAGCTCCTGGGATT (SEQ ID NO:137);  
GGCTGGATAGCAATT (SEQ ID NO:138); CTAGCTCCTGGGAAT (SEQ ID NO:139);

40 CACCTGGCTCTGCCC (SEQ ID NO:140); GGGACTTGCCTAGGG (SEQ ID NO:141);  
CACCTGGCTCTGCAC (SEQ ID NO:142); GGGACTTGCCTAGTG (SEQ ID NO:143);

CCTGGCCCCCACC (SEQ ID NO:144); CTCTGATGTGAGGGG (SEQ ID NO:145);  
45 CCTGGCCCCCACTC (SEQ ID NO:146); CTCTGATGTGAGGAG (SEQ ID NO:147);

GAACCCTCCCTCCGC (SEQ ID NO:148); GCTGGCTGCGATGCG (SEQ ID NO:149);  
GAACCCTCCCTCCAC (SEQ ID NO:150); GCTGGCTGCGATGTG (SEQ ID NO:151);

50 TAGATACACCTGCTG (SEQ ID NO:152); GCAGATACACCACAG (SEQ ID NO:153);  
TAGATACACCTGCCG (SEQ ID NO:154); GCAGATACACCACGG (SEQ ID NO:155);

GAAGGCATGTCTGAA (SEQ ID NO:156); ATGGCTGTCTACTTC (SEQ ID NO:157);  
GAAGGCATGTCTGCA (SEQ ID NO:158); ATGGCTGTCTACTGC (SEQ ID NO:159);

55 GAACCCTGACCAACC (SEQ ID NO:160); TGCAAAAGCAAAGGA (SEQ ID NO:161);  
GAACCCTGACCAATC (SEQ ID NO:162); TGCAAAAGCAAAGAA (SEQ ID NO:163);

TCTTGCCCTGTTTTTC (SEQ ID NO:164); TGTTGTGCTCCAGAA (SEQ ID NO:165);  
 TCTTGCCCTGTTTTC (SEQ ID NO:166); TGTTGTGCTCCAGGA (SEQ ID NO:167);  
 5 TGTTCCCTGGACCTGC (SEQ ID NO:168); TCTCCTCTCCGAGCA (SEQ ID NO:169);  
 TGTTCCCTGGACCTTC (SEQ ID NO:170); TCTCCTCTCCGAGAA (SEQ ID NO:171);  
 TGGGGGAGTCATGCC (SEQ ID NO:172); AAGGTGGAAGAAGGC (SEQ ID NO:173);  
 TGGGGGAGTCATGTC (SEQ ID NO:174); AAGGTGGAAGAAGAC (SEQ ID NO:175);  
 10 AGTCATGCCTTCTTC (SEQ ID NO:176); TTCCCGAAGGTGGAA (SEQ ID NO:177);  
 AGTCATGCCTTCTCC (SEQ ID NO:178); TTCCCGAAGGTGGGA (SEQ ID NO:179);  
 CAGCTGCAGCCCCCG (SEQ ID NO:180); TGGGGGCCGAGACGG (SEQ ID NO:181);  
 15 CAGCTGCAGCCCCTG (SEQ ID NO:182); TGGGGGCCGAGACAG (SEQ ID NO:183);  
 CTATCAGGAGTTTGT (SEQ ID NO:184); TCCACCGCATGTACA (SEQ ID NO:185);  
 CTATCAGGAGTTTAT (SEQ ID NO:186); TCCACCGCATGTATA (SEQ ID NO:187);  
 20 CCCAAGCAGCTCCCC (SEQ ID NO:188); CCCAGGTGCTCTGGG (SEQ ID NO:189);  
 CCCAAGCAGCTCCTC (SEQ ID NO:190); CCCAGGTGCTCTGAG (SEQ ID NO:191);  
 CTGTGGAGACAGGTC (SEQ ID NO:192); GTAGGGGGCGAGGAC (SEQ ID NO:193);  
 CTGTGGAGACAGGGC (SEQ ID NO:194); GTAGGGGGCGAGGCC (SEQ ID NO:195);  
 25 TCCATCCTGCCCCCTG (SEQ ID NO:196); TCTGAGCATTGCCAG (SEQ ID NO:197);  
 TCCATCCTGCCCCCG (SEQ ID NO:198); TCTGAGCATTGCCGG (SEQ ID NO:199);  
 TCTCTTAGGTGCATG (SEQ ID NO:200); GCAACAAGAGGACAT (SEQ ID NO:201);  
 30 TCTCTTAGGTGCACG (SEQ ID NO:202); GCAACAAGAGGACGT (SEQ ID NO:203);  
 TCTTAGGTGCATGTC (SEQ ID NO:204); CAGCAACAAGAGGAC (SEQ ID NO:205);  
 TCTTAGGTGCATGCC (SEQ ID NO:206); CAGCAACAAGAGGGC (SEQ ID NO:207);  
 GACTAGGGCTTATCC (SEQ ID NO:208); TTTCCCAGGCATGGA (SEQ ID NO:209);  
 35 GACTAGGGCTTATTC (SEQ ID NO:210); TTTCCCAGGCATGAA (SEQ ID NO:211);  
 GAAGGCAGCCAGGCT (SEQ ID NO:212); TGGAAATCTGCCAGC (SEQ ID NO:213);  
 GAAGGCAGCCAGGGT (SEQ ID NO:214); TGGAAATCTGCCACC (SEQ ID NO:215);  
 40 GATCATGGCCACAT (SEQ ID NO:216); AGGTGGGCCCTCCATG (SEQ ID NO:217);  
 GATCATGGCCACGT (SEQ ID NO:218); AGGTGGGCCCTCCACG (SEQ ID NO:219);  
 AGAAACTAACACAGC (SEQ ID NO:220); ATTCCCTTGATGGCT (SEQ ID NO:221);  
 45 AGAAACTAACACAAC (SEQ ID NO:222); ATTCCCTTGATGGTT (SEQ ID NO:223);  
 GTTGAGTAATGCTCG (SEQ ID NO:224); AAAACACACAGACGA (SEQ ID NO:225);  
 GTTGAGTAATGCTTG (SEQ ID NO:226); AAAACACACAGACAA (SEQ ID NO:227);  
 TAAGAAACTTGCCGT (SEQ ID NO:228); ACCCAAACCCAGACG (SEQ ID NO:229);  
 50 TAAGAAACTTGCCAT (SEQ ID NO:230); and ACCCAAACCCAGATG (SEQ ID NO:231).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the

novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3'-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site. A particularly preferred oligonucleotide primer for detecting IL4R $\alpha$

5. gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

	AGATTGCACC (SEQ ID NO:232);	TGGAGTGCAG (SEQ ID NO:233);
	TGCTTTTGTG (SEQ ID NO:234);	AAGGGGAATA (SEQ ID NO:235);
	TTCCTGGGCC (SEQ ID NO:236);	GCAGCCTGAG (SEQ ID NO:237);
10	GAGTAAGCCT (SEQ ID NO:238);	AGCTCCAGCG (SEQ ID NO:239);
	CTGAGAACAA (SEQ ID NO:240);	CCGCGCCTCC (SEQ ID NO:241);
	CAACGGAGGC (SEQ ID NO:242);	ACGCACCCCG (SEQ ID NO:243);
	TCAGTGC GGA (SEQ ID NO:244);	GTGTATAGTT (SEQ ID NO:245);
	GGGCGGAGTG (SEQ ID NO:246);	CACCCCTGCC (SEQ ID NO:247);
15	GCTGCCTGGG (SEQ ID NO:248);	CCCACCCTCA (SEQ ID NO:249);
	GGGTGGGGTG (SEQ ID NO:250);	CCTCCCCTGC (SEQ ID NO:251);
	CGCTTCTCCC (SEQ ID NO:252);	GGTTTCACTG (SEQ ID NO:253);
	GGAGTGAAAA (SEQ ID NO:254);	CTGCCGGGTC (SEQ ID NO:255);
	GGAGGCAAGC (SEQ ID NO:256);	CCAGCCCCAG (SEQ ID NO:257);
20	CAAGCCCTGG (SEQ ID NO:258);	GCTATCCAGC (SEQ ID NO:259);
	TGGATAGCAA (SEQ ID NO:260);	GCTCCTGGGA (SEQ ID NO:261);
	CTGGCTCTGC (SEQ ID NO:262);	ACTTGCCTAG (SEQ ID NO:263);
	GGCCCCCACC (SEQ ID NO:264);	TGATGTGAGG (SEQ ID NO:265);
	CCCTCCCTCC (SEQ ID NO:266);	GGCTGCGATG (SEQ ID NO:267);
25	ATACACCTGC (SEQ ID NO:268);	GATACACCAC (SEQ ID NO:269);
	GGCATGTCTG (SEQ ID NO:270);	GCTGTCTACT (SEQ ID NO:271);
	CCCTGACCAA (SEQ ID NO:272);	AAAAGCAAAG (SEQ ID NO:273);
	TGCCCTGTTT (SEQ ID NO:274);	TGTGCTCCAG (SEQ ID NO:275);
	TCCTGGACCT (SEQ ID NO:276);	CCTCTCCGAG (SEQ ID NO:277);
30	GGGAGTCATG (SEQ ID NO:278);	GTGGAAGAAG (SEQ ID NO:279);
	CATGCCTTCT (SEQ ID NO:280);	CCGAAGGTGG (SEQ ID NO:281);
	CTGCAGCCCC (SEQ ID NO:282);	GGGCCGAGAC (SEQ ID NO:283);
	TCAGGAGTTT (SEQ ID NO:284);	ACCGCATGTA (SEQ ID NO:285);
	AAGCAGCTCC (SEQ ID NO:286);	AGGTGCTCTG (SEQ ID NO:287);
35	TGGAGACAGG (SEQ ID NO:288);	GGGGGCGAGG (SEQ ID NO:289);
	ATCCTGCCCC (SEQ ID NO:290);	GAGCATTGCC (SEQ ID NO:291);
	CTTAGGTGCA (SEQ ID NO:292);	ACAAGAGGAC (SEQ ID NO:293);

TAGGTGCATG (SEQ ID NO:294); CAACAAGAGG (SEQ ID NO:295);  
 TAGGGCTTAT (SEQ ID NO:296); CCCAGGCATG (SEQ ID NO:297);  
 GGCAGCCAGG (SEQ ID NO:298); AAATCTGCCA (SEQ ID NO:299);  
 CATGGCCAC (SEQ ID NO:300); TGGGCCTCCA (SEQ ID NO:301);  
 5 AACTAACACA (SEQ ID NO:302); CCCTTGATGG (SEQ ID NO:303);  
 GAGTAATGCT (SEQ ID NO:304); ACACACAGAC (SEQ ID NO:305);  
 GAAACTTGCC (SEQ ID NO:306); and CAAACCCAGA (SEQ ID NO:307).

In some embodiments, a composition contains two or more differently labeled genotyping  
 10 oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites.  
 It is also contemplated that primer compositions may contain two or more sets of allele-specific primer  
 pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic  
 site.

IL4R $\alpha$  genotyping oligonucleotides of the invention may also be immobilized on or synthesized  
 15 on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019).  
 Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection  
 assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized  
 IL4R $\alpha$  genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides  
 designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

20 In another embodiment, the invention provides a kit comprising at least two genotyping  
 oligonucleotides packaged in separate containers. The kit may also contain other components such as  
 hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate  
 container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may  
 contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer  
 25 extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping  
 and/or haplotyping the IL4R $\alpha$  gene in an individual. As used herein, the terms "IL4R $\alpha$  genotype" and  
 "IL4R $\alpha$  haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide,  
 respectively, that is present at one or more of the novel polymorphic sites described herein and may  
 30 optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic  
 sites in the IL4R $\alpha$  gene. The additional polymorphic sites may be currently known polymorphic sites or  
 sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid  
 mixture comprising the two copies of the IL4R $\alpha$  gene, or a fragment thereof, that are present in the  
 35 individual, and determining the identity of the nucleotide pair at one or more of the polymorphic sites  
 selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16,  
 PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36.

PS37, PS38, PS39, PS40, PS41, PS44, and PS45 in the two copies to assign an IL4R $\alpha$  genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a preferred embodiment of the genotyping method, the identity of the nucleotide pair at one or more of the polymorphic sites selected from the group consisting of PS5, PS24, PS26, PS29, PS31, PS42, and PS43 is also determined. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-45.

Typically, the nucleic acid mixture is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid mixture may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from an organ in which the IL4R $\alpha$  gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5' and 3' nontranscribed regions. If an IL4R $\alpha$  gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid molecule containing only one of the two copies of the IL4R $\alpha$  gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more of the polymorphic sites PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45 in that copy to assign an IL4R $\alpha$  haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the IL4R $\alpha$  gene or fragment such as one of the methods described above for preparing IL4R $\alpha$  isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two IL4R $\alpha$  gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional IL4R $\alpha$  clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the IL4R $\alpha$  gene in an individual. In some embodiments, the haplotyping method also comprises identifying the nucleotide at one or more of the polymorphic sites PS5, PS24, PS26, PS29, PS31, PS42, and PS43. In a particularly preferred embodiment, the nucleotide at each of PS1-45 is identified.

In a preferred embodiment, an IL4R $\alpha$  haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more of the polymorphic sites selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45 in each copy of the IL4R $\alpha$  gene that is present in the individual. In a

particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-45 in each copy of the IL4R $\alpha$  gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the IL4R $\alpha$  gene, or fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping a polymorphic site not disclosed herein that is in linkage disequilibrium with the polymorphic site that is of interest. Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stevens, JC 1999, *Mol. Diag.* 4: 309-17). Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., *Proc. Natl. Acad. Sci. USA* 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landgren et al., *Science* 241:1077-1080, 1988). Oligonucleotides useful as primers or probes in such methods should specifically hybridize to a region of the nucleic acid

that contains or is adjacent to the polymorphic site. Typically, the oligonucleotides are between 10 and 35 nucleotides in length and preferably, between 15 and 30 nucleotides in length. Most preferably, the oligonucleotides are 20 to 25 nucleotides long. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan.

5 Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* 89:392-396, 1992).

10 A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or  
15 oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or  
20 the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include  
25 substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the IL4R $\alpha$  gene of an individual may also be determined by  
30 hybridization of a nucleic sample containing one or both copies of the gene to nucleic acid arrays and subarrays such as described in WO 95/1 1995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., *Proc. Natl.*  
35 *Acad. Sci. USA* 82:7575, 1985; Meyers et al., *Science* 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the *E. coli* mutS protein (Modrich, *P. Ann. Rev. Genet.* 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism

(SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

5 A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism  
10 may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruano et al., Nucl. Acids Res. 17:8392, 1989; Ruano et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al.  
15 (WO89/10414).

In another aspect of the invention, an individual's IL4R $\alpha$  haplotype pair is predicted from its IL4R $\alpha$  genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying an IL4R $\alpha$  genotype for the individual at two or more polymorphic sites selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9,  
20 PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing IL4R $\alpha$  haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include  
25 the IL4R $\alpha$  haplotype pairs shown in Table 4.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number  
30 of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by  $2n = \log(1-q)/\log(1-p)$  where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99%  
35 certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.



In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3<sup>rd</sup> Ed., 1997) postulates that the frequency of finding the haplotype pair  $H_1 / H_2$  is equal to

5  $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$  if  $H_1 \neq H_2$  and  $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$  if  $H_1 = H_2$ . A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group  
10 can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

15 In one embodiment of this method for predicting an IL4Rα haplotype pair, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype  
20 pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular  
25 haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

The invention also provides a method for determining the frequency of an IL4Rα genotype or IL4Rα haplotype in a population. The method comprises determining the genotype or the haplotype pair  
30 for the IL4Rα gene that is present in each member of the population, wherein the genotype or haplotype comprises the nucleotide pair or nucleotide detected at one or more of the polymorphic sites PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45 in the IL4Rα gene; and calculating the frequency any particular genotype or  
35 haplotype is found in the population. The population may be a reference population, a family population, a same sex population, a population group, a trait population (e.g., a group of individuals exhibiting a trait

of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for IL4R $\alpha$  genotypes and/or haplotypes found in a reference population are used in a method for identifying an association between a trait and an IL4R $\alpha$  genotype or an IL4R $\alpha$  haplotype. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s) or haplotype(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s) or haplotype(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes and/or haplotypes observed in the populations are compared. If a particular genotype or haplotype for the IL4R $\alpha$  gene is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that IL4R $\alpha$  genotype or haplotype. Preferably, the IL4R $\alpha$  genotype or haplotype being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 4 and 5, respectively, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting IL4R $\alpha$  or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and an IL4R $\alpha$  genotype or haplotype, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the IL4R $\alpha$  gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and IL4R $\alpha$  genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their IL4R $\alpha$  genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the IL4R $\alpha$  gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention starts with a model of the form

$$r = r_0 + S \times d$$

where  $r$  is the response,  $r_0$  is a constant called the "intercept",  $S$  is the slope and  $d$  is the dose. To determine the dose, the most-common and least common nucleotides at the polymorphic site are first defined. Then, for each individual in the trial population, one calculates a "dose" as the number of least-common nucleotides the individual has at the polymorphic site of interest. This value can be 0 (homozygous for the least-common nucleotide), 1 (heterozygous), or 2 (homozygous for the most common nucleotide). An individual's "response" is the value of the clinical measurement. Standard linear regression methods are then used to fit all the individuals' doses and responses to a single model (see e.g., L.D. Fisher and G. vanBelle, *supra*, Ch 9). The outputs of the regression calculation are the

intercept  $r_0$ , the slope  $S$ , and the variance (which measures how well the data fits this simple linear model). The Students t-test value and the level of significance can then be calculated for each of the polymorphic sites.

A second method for finding correlations between IL4R $\alpha$  haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in *Reviews in Computational Chemistry*, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2<sup>nd</sup> Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., *supra* Ch. 10), or other global or local optimization approaches (see discussion in Judson, *supra*) could also be used. As an example, a genetic algorithm approach is described herein. This method searches for optimal parameters or weights in linear or non-linear models connecting IL4R $\alpha$  haplotype loci and clinical outcome. One model is of the form

$$C = C_0 + \sum_{\alpha} \left( \sum_i w_{i,\alpha} R_{i,\alpha} + \sum_i w'_{i,\alpha} L_{i,\alpha} \right) \quad [1]$$

where  $C$  is the measured clinical outcome,  $i$  goes over all polymorphic sites,  $\alpha$  over all candidate genes,  $C_0$ ,  $w_{i,\alpha}$  and  $w'_{i,\alpha}$  are variable weight values,  $R_{i,\alpha}$  is equal to 1 if site  $i$  in gene  $\alpha$  in the first haplotype takes on the most common nucleotide and  $-1$  if it takes on the less common nucleotide.  $L_{i,\alpha}$  is the same as  $R_{i,\alpha}$  except for the second haplotype. The constant term  $C_0$  and the weights  $w_{i,\alpha}$  and  $w'_{i,\alpha}$  are varied by the genetic algorithm during a search process that minimizes the error between the measured value of  $C$  and the value calculated from Equation 1. Models other than the one given in Equation 1 can be readily incorporated by those skilled in the art for analyzing the clinical and polymorphism data. The genetic algorithm is especially suited for searching not only over the space of weights in a particular model but also over the space of possible models (Judson, *supra*).

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the IL4R $\alpha$  gene. ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, *supra*, Ch. 10). These traits or variables are called the independent variables. To carry out ANOVA, the independent variable(s) are measured and individuals are placed into groups based on their values for these variables. In this case, the independent variable(s) refers to the combination of polymorphisms present at a subset of the polymorphic sites, and thus, each group contains those individuals with a given genotype or haplotype pair. The variation in response within the groups and also the variation between groups is then measured. If the within-group response variation is large (people in a group have a wide range of responses) and the

response variation between groups is small (the average responses for all groups are about the same) then it can be concluded that the independent variables used for the grouping are not causing or correlated with the response variable. For instance, if people are grouped by month of birth (which should have nothing to do with their response to a drug) the ANOVA calculation should show a low level of significance.

5 However, if the response variation is larger between groups than within groups, the F-ratio ("between groups" divided by "within groups") is greater than one. Large values of the F-ratio indicate that the independent variable is causing or correlated with the response. The calculated F-ratio is preferably compared with the critical F-distribution value at whatever level of significance is of interest. If the F-ratio is greater than the Critical F-distribution value, then one may be confident that the individual's  
10 genotype or haplotype pair for this particular subset of polymorphic sites in the IL4R $\alpha$  gene is at least partially responsible for, or is at least strongly correlated with the clinical response.

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of IL4R $\alpha$  genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

15 The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the IL4R $\alpha$  gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the  
20 polymorphic sites in the IL4R $\alpha$  gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying IL4R $\alpha$  genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

Any or all analytical and mathematical operations involved in practicing the methods of the  
25 present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the IL4R $\alpha$  gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin,  
30 clinical responses, genotypes, and haplotypes for one or more populations). The IL4R $\alpha$  polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the  
35 computer via a network.

Preferred embodiments of the invention are described in the following examples. Other

embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

5

### EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2<sup>nd</sup> Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

10

15

#### Example 1A

This example illustrates examination of various regions of the IL4R $\alpha$  gene for polymorphic sites using DNA from Index Repository IA.

#### Amplification of Target Regions

20

The following target regions of the IL4R $\alpha$  gene were amplified using the PCR primer pairs listed below, with the sequences presented in the 5' to 3' direction and nucleotide positions shown for each region corresponding to the indicated GenBank Accession No.

Accession Number: AC004525

Fragment 1

25

Forward Primer

32801-32822 CCACAGTCATCCCGACACTAGC (SEQ ID NO:308)

Reverse Primer

Complement of 33355-33334 TATCCAGCCGTATCCATGTGC (SEQ ID NO:309)

PCR product 555 nt

30

Fragment 2

Forward Primer

35515-35536 CCTTGGTGCATGTGGTAAGAGG (SEQ ID NO:310)

Reverse Primer

35

Complement of 36068-36046 TTTCAAAGGTGGGAGGACTGAGG (SEQ ID NO:311)

PCR product 554 nt

Fragment 3

Forward Primer

40

37031-37050 GCAGTGAGCTGGGATTGTGC (SEQ ID NO:312)

Reverse Primer

Complement of 37701-37679 AACTCCCCTTCTCTGATGTGAGG (SEQ ID NO:313)

PCR product 671 nt

## Fragment 4

## Forward Primer

43240-43262 TCACAGTTACAGAGGTGGCAAGC (SEQ ID NO:314)

## Reverse Primer

5 Complement of 43727-43706 CTGCCTACCTGGCAGATACACC (SEQ ID NO:315)

PCR product 488 nt

## Fragment 5

## Forward Primer

10 49553-49574 AGCTGTCACCTCCACCTCCTTGG (SEQ ID NO:316)

## Reverse Primer

Complement of 50036-50013 AAAGCCTCTGGTCTGCTAATGACC (SEQ ID NO:317)

PCR product 484 nt

## 15 Fragment 6

## Forward Primer

51394-51415 GGGAGGAGATTCAGAGCACTCC (SEQ ID NO:318)

## Reverse Primer

Complement of 51847-51826 CAGTCCACGTTTCCAGAACACC (SEQ ID NO:319)

20 PCR product 454 nt

## Fragment 7

## Forward Primer

52806-52826 GGCTTGGGATAATGGTGTTC (SEQ ID NO:320)

## 25 Reverse Primer

Complement of 53529-53507 TACTTCCCGAAGGTGGAAGAAGG (SEQ ID NO:321)

PCR product 724 nt

## Fragment 8

## 30 Forward Primer

53242-53265 CAGTGGAGATCAGCAAGACAGTCC (SEQ ID NO:322)

## Reverse Primer

Complement of 53807-53786 GGGCATCTCGGGTTCTACTTCC (SEQ ID NO:323)

PCR product 566 nt

35

## Fragment 9

## Forward Primer

53522-53544 GGGAAGTACGAGTGCTCACATGC (SEQ ID NO:324)

## Reverse Primer

40 Complement of 54110-54088 CTTATACCCCTCTTCCCCACTGC (SEQ ID NO:325)

PCR product 589 nt

## Fragment 10

## Forward Primer

45 53821-53841 TCTCTGAGCCAACCACTGTGC (SEQ ID NO:326)

## Reverse Primer

Complement of 54359-54337 GGCTGAGTAGACAATGCCACTGC (SEQ ID NO:327)

PCR product 539 nt

## 50 Fragment 11

## Forward Primer

54055-54076 CTGTGTCCCCAGAGAAATGTGG (SEQ ID NO:328)

## Reverse Primer

Complement of 54717-54695 GACTCAGCAACAAGAGGACATGC (SEQ ID NO:329)

55 PCR product 663 nt

## Fragment 12

## Forward Primer

54342-54365 GGCATTGTCTACTCAGCCCTTACC (SEQ ID NO:330)

## Reverse Primer

5 Complement of 54986-54967 ACAAGTCGAGGTGCCCAAGG (SEQ ID NO:331)

PCR product 645 nt

## Fragment 13

## Forward Primer

10 54669-54693 CCCACATACATGAGGGTCTCTTAGG (SEQ ID NO:332)

## Reverse Primer

Complement of 55270-55250 ATTCTGCCTCCAGCATCAACC (SEQ ID NO:333)

PCR product 602 nt

## 15 Fragment 14

## Forward Primer

55235-55258 AACAGAGCTTCCTTAGGTTGATGC (SEQ ID NO:334)

## Reverse Primer

Complement of 55847-55825 CCTCAGTTCCCCACTACCTTAGC (SEQ ID NO:335)

20 PCR product 613 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of Index Repository IA. The PCR reactions were carried out under the following conditions:

25	Reaction volume	= 20 µl
	10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
	100 ng of human genomic DNA	= 1 µl
	10 mM dNTP	= 0.4 µl
	Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
30	Forward Primer (10 µM)	= 0.4 µl
	Reverse Primer (10 µM)	= 0.4 µl
	Water	= 15.6 µl

## Amplification profile:

35	94°C - 2 min.	1 cycle
	94°C - 30 sec.	} 10 cycles
	70°C - 45 sec.	
	72°C - 1 min.	
40	94°C - 30 sec.	} 35 cycles
	64°C - 45 sec.	
	72°C - 1 min.	

45 Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at [http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI\\_pcr.html](http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html).

Briefly, five µl of carboxyl coated magnetic beads (10 mg/ml) and 60 µl of HYB BUFFER (2.5M



NaCl/20% PEG 8000) were added to each PCR reaction mixture (20 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25 µl of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously or those listed, in the 5' to 3' direction, below.

Accession Number: AC004525

Fragment 1

Forward Primer

32865-32882 GCGCTGGCCCTCAACTTT (SEQ ID NO:336)

Reverse Primer

Complement of 33283-33264 GTCCCTGGAGATGGGACCTC (SEQ ID NO:337)

Fragment 2

Forward Primer

35598-35617 GCCCCAGATCTGTCTCAC (SEQ ID NO:338)

Reverse Primer

Complement of 36013-35994 GGAAAATACAGGCGGCTTCC (SEQ ID NO:339)

Fragment 3

Forward Primer

37182-37203 GGCTCTGAATCTGTGTGGTGCT (SEQ ID NO:340)

Reverse Primer

Complement of 37639-37620 AGCCAGGTGAGAAGCCAGGT (SEQ ID NO:341)

Fragment 4

Forward Primer

43266-43285 GGCCTGAACAGGACGAACAA (SEQ ID NO:342)

Reverse Primer

Complement of 43687-43668 GGCAGGATTGCCATTAGAGG (SEQ ID NO:343)

Fragment 5

Forward Primer

49639-49660 TGAGTCAGTGGTTTGACCTCCA (SEQ ID NO:344)

Reverse Primer

Complement of 49999-49980 GCCTCTGTCTCCCCTGCAAC (SEQ ID NO:345)

Fragment 6

Forward Primer

51423-51442 CCACCTTTGCCATCGACCAC (SEQ ID NO:346)

Reverse Primer

Complement of 51818-51799 CTGCCGTCCCTGAAGGCTA (SEQ ID NO:347)

Fragment 7

Forward Primer

52932-52953 CCCTACCCTCAGGGATTTCTCA (SEQ ID NO:348)

Reverse Primer

Complement of 53474-53455 CCCATTCTCTCTCCGAGCA (SEQ ID NO:349)

## Fragment 8

## Forward Primer

53280-53299 ATCAGCGTGGTGCGATGTGT (SEQ ID NO:350)

## Reverse Primer

5 Complement of 53756-53737 ACCCAGCTCTCTGGGACACG (SEQ ID NO:351)

## Fragment 9

## Forward Primer

53548-53568 GGGATGAGTTCCCAAGTGCAG (SEQ ID NO:352)

## 10 Reverse Primer

Complement of 54048-54029 TGGCAAGCAGGCTTGAGAAG (SEQ ID NO:353)

## Fragment 10

## Forward Primer

15 53841-53860 CCCCAACCTGAGCCAGAAAC (SEQ ID NO:354)

## Reverse Primer

Complement of 54330-54311 TGTCCACAAGGGGGTCTGTG (SEQ ID NO:355)

## Fragment 11

## 20 Forward Primer

54083-54102 GGCTAGCAGTGGGGAAGAGG (SEQ ID NO:356)

## Reverse Primer

Complement of 54617-54600 ATTGCCAGGGGCAGGATG (SEQ ID NO:357)

## 25 Fragment 12

## Forward Primer

54422-54441 CCCTGTCATGGCCAGTCCTT (SEQ ID NO:358)

## Reverse Primer

30 Complement of 54929-54910 GCGACCCAGTGCCCTCTACT (SEQ ID NO:359)

## Fragment 13

## Forward Primer

54694-54715 TGCATGTCCTCTTGTTGCTGAG (SEQ ID NO:360)

## Reverse Primer

35 Complement of 55213-55194 CAATGACCACCCTCCCTGAA (SEQ ID NO:361)

## Fragment 14

## Forward Primer

55273-55291 CGGCTGTCAAGGGGTGTTC (SEQ ID NO:362)

## 40 Reverse Primer

Complement of 55769-55750 CCAAACCCAGACGGCAAGTT (SEQ ID NO:363)

45 Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., Nucleic Acids Res. 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the IL4R $\alpha$  gene are listed in Table 3 below.

50

Example 1B

This example illustrates examination of the IL4R $\alpha$  gene for polymorphic sites in the following

target regions: 1000 base pairs upstream of the ATG start codon; each of the exons, including approximately 100 base pairs on either side of the exon; and 500-1000 base pairs downstream of the termination codon.

## 5 Amplification of the Target Regions

PCR primers, which were designed based on the nearly complete IL4R genomic sequence reported in the GenBank database (Accession No: AC004525), are set forth below:

### Promoter

#### Forward Primer

10 30094-30117 AAACCCCTGGACTCCAAGTGATCC (SEQ ID NO:364)

#### Reverse Primer

Complement of 30842-30819 AAGCGATTCTTCTGCCTCAGCCTC (SEQ ID NO:365)

PCR product 749 nt

### 15 Exon 1

#### Forward Primer

30517-30540 GGACAGTTGTTGTGTAGCTCACCC (SEQ ID NO:366)

#### Reverse Primer

Complement of 31409-31431 CTATGTTGCCCAAGCTGACCTC (SEQ ID NO:367)

20 PCR product 893 nt

### Exon 2

#### Forward Primer

32871-32890 GCCCTCAACTTTGCCTGCAC (SEQ ID NO:368)

#### 25 Reverse Primer

Complement of 33362-33340 AGTCCAGTATTCCAGCCGTATCC (SEQ ID NO:369)

PCR product 492 nt

### Exon 3

#### 30 Forward Primer

35623-35644 TGATCGGGAAGCTGGAAGAGTC (SEQ ID NO:370)

#### Reverse Primer

Complement of 36070-36049 CGTTTCAAAGGTGGGAGGACTG (SEQ ID NO:371)

PCR product 448 nt

35

### Exon 4

#### Forward Primer

37230-37250 CGACCAAAAATCTGGGTGGTG (SEQ ID NO:372)

#### Reverse Primer

40 Complement of 37668-37648 CAGGAAGCAAAGGGACTTGCC (SEQ ID NO:373)

PCR product 439 nt

### Exon 5

#### Forward Primer

45 43307-43328 TCTTAAACATGGTGGGGTCAGC (SEQ ID NO:374)

#### Reverse Primer

Complement of 43765-43745 CATGGAAATTGTGGGCTTGTG (SEQ ID NO:375)

PCR product 459 nt

### 50 Exon 6

#### Forward Primer

46507-46528 ATGTGCAAGAGGGAGAGTGGTG (SEQ ID NO:376)

## Reverse Primer

Complement of 46870-46849 TGACTGAGAGGACTGCAAAGGG (SEQ ID NO:377)

PCR product 364 nt

## 5 Exon 7

## Forward Primer

49672-49695 GCCTGATCTCTGATGCCAAATAAG (SEQ ID NO:378)

## Reverse Primer

Complement of 49964-49983 TTTGCCATTCCAGAAGCCAG (SEQ ID NO:379)

10 PCR product 293 nt

## Exon 8

## Forward Primer

51524-51546 GATCTGTGTGATGTCGAGGCTTG (SEQ ID NO:380)

## 15 Reverse Primer

Complement of 51845-51825 GTCCACGTTTCCAGAACACCC (SEQ ID NO:381)

PCR product 322 nt

## Exon 9 Fragment 1

## 20 Forward Primer

52912-52933 CGAAATCCCAAAGACACAGACC (SEQ ID NO:382)

## Reverse Primer

Complement of 53722-53701 GAGTTGCTGAAGCTGCGGTAAG (SEQ ID NO:383)

PCR product 811 nt

25

## Exon 9 Fragment 2

## Forward Primer

53352-53374 GAAAAAGGGAGCTTCTGTGCATC (SEQ ID NO:384)

## Reverse Primer

30 Complement of 54172-54153 AACAAAGGGGACAGGGACTGG (SEQ ID NO:385)

PCR product 821 nt

## Exon 9 Fragment 3

## Forward Primer

35 53854-53874 CAGAAACCTGGGAGCAGATCC (SEQ ID NO:386)

## Reverse Primer

Complement of 54711-54688 GCAACAAGAGGACATGCACCTAAG (SEQ ID NO:387)

PCR product 858 nt

## 40 Exon 9 Fragment 4

## Forward Primer

54266-54288 AAAGGTAGAGGACATGCCAAAGC (SEQ ID NO:388)

## Reverse Primer

Complement of 55007-54987 GGAGCAGCCAACAACCTCGTTC (SEQ ID NO:389)

45 PCR product 742 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for a reference population of 70 human individuals. The PCR reactions were carried out under the following conditions:

50	Reaction volume	= 20 $\mu$ l
	10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 $\mu$ l
	100 ng of human genomic DNA	= 1 $\mu$ l
	10 mM dNTP	= 0.4 $\mu$ l

Advantage 2 Polymerase enzyme mix (Clontech)

= 0.2  $\mu$ l

Forward Primer (10  $\mu$ M)

= 0.4  $\mu$ l

Reverse Primer (10  $\mu$ M)

= 0.4  $\mu$ l

Water

=15.6 $\mu$ l

5

Amplification profile:

94°C - 2 min. 1 cycle

10

94°C - 30 sec.

70°C - 45 sec.

72°C - 1 min.

} 10 cycles

15

94°C - 30 sec.

64°C - 45 sec.

72°C - 1 min.

} 35 cycles

#### Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at [http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI\\_pcr.html](http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html).

Briefly, five  $\mu$ l of carboxyl coated magnetic beads (10 mg/ml) and 60  $\mu$ l of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (20  $\mu$ l). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150  $\mu$ l of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25  $\mu$ l of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously or those listed, in the 5' to 3' direction, below.

30

Promoter Fragment 1

Forward Primer

30312-30334 GCTCATTTAATCCCCACAACACC (SEQ ID NO:390)

Reverse Primer

35 Complement of 30791-30769 CCACCACACCTGGCTAATTTTGT (SEQ ID NO:391)

Promoter Fragment 2

Forward Primer

30529-30551 TGTAGCTCACCCCTCTGGACTTTG (SEQ ID NO:392)

40

Reverse Primer

Complement of 30990-30971 AATATGCAACCCTCCCCTGC (SEQ ID NO:393)

Exon 1

Forward Primer

45 30824-30846 TGAGGCAGAAGAATCGCTTGAAC (SEQ ID NO:394)

Reverse Primer

Complement of 31261-31240 ACTTGTCATTGGCTGTCCCCTC (SEQ ID NO:395)

## Exon 2

## Forward Primer

32880-32900 TTTGCCTGCACTGTGCTTTTG (SEQ ID NO:396)

## Reverse Primer

Complement of 33234-33212 CCATACTCAGCATCCTGCACTCC (SEQ ID NO:397)

## Exon 5

## Forward Primer

43330-43349 AACGACAGCAACCAGGGTGG (SEQ ID NO:398)

## Reverse Primer

Complement of 43704-43682 CAGCAGGTGTATCTAATGGCAGG (SEQ ID NO:399)

## Exon 6

## Forward Primer

46520-46541 AGAGTGGTGGGGAGATGAGGTG (SEQ ID NO:400)

## Reverse Primer

Complement of 46857-46837 TGCAAAGGGGCAGACTAGAGG (SEQ ID NO:401)

## Exon 7

## Forward Primer

49708-49729 CGACCACTTTTATGGGAGGAGC (SEQ ID NO:402)

## Reverse Primer

Complement of 49927-49905 CCAGGTGTTCTGAACCACACTTC (SEQ ID NO:403)

## Exon 8

## Forward Primer

51528-51550 TGTGTGATGTCGAGGCTTGTACC (SEQ ID NO:404)

## Reverse Primer

Complement of 51779-51758 GAATGCAGGGAAGAGAAGGCAG (SEQ ID NO:405)

## Exon 9 Fragment 1

## Forward Primer

53017-53038 GCCATCAGGACATGGTGATTTC (SEQ ID NO:406)

## Reverse Primer

Complement of 53539-53518 TGAGCACTCGTACTTCCCGAAG (SEQ ID NO:407)

## Exon 9 Fragment 2

## Forward Primer

53378-53399 TGAGAGCAGCAGGGATGACTTC (SEQ ID NO:408)

## Reverse Primer

Complement of 53948-53926 AAATCCTGATAGCCACTGGTGG (SEQ ID NO:409)

## Exon 9 Fragment 3

## Forward Primer

53869-53868 AGATCCTCCGCCGAAATGTC (SEQ ID NO:410)

## Reverse Primer

Complement of 54583-54560 TTACTCTTCTCTGAGATGCCCGAG (SEQ ID NO:411)

## Exon 9 Fragment 4

## Forward Primer

54334-54355 TGGGCAGTGGCATTGTCTACTC (SEQ ID NO:412)

## Reverse Primer

Complement of 54769-54750 TTCCAGGAGGTGGCATTTC (SEQ ID NO:413)

Sequencing reactions were performed using the Big-Dye terminator kit from PE Biosystems (Foster City, CA) according to the manufacturer's instructions. The sequencing products were analyzed on an ABI 477 automated sequencer (PE Biosystems, Foster City, CA).

5 Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., 14 *Nucleic Acids Res.* 2745-2751, 1997). The presence of a polymorphism was confirmed on both the strands. The polymorphisms and their locations in the IL4R $\alpha$  gene are listed in Table 3 below.

Table 3. Polymorphic Sites Identified in the IL4R $\alpha$  Gene

Polymorphic Site Number	Nucleotide Position in GenBank Accession	Nucleotide Position in Figure 1	Reference Allele	Variant Allele	Example
PS1	97137(Acc#AC004525)	32884	A	G	1B
PS2	97118(Acc#AC004525)	32903	C	T	1A
PS3	97060(Acc#AC004525)	32961	G	T	1A
PS4	96886(Acc#AC004525)	33135	G	C	1A
PS5 <sup>R</sup>	94272(Acc#AC004525)	35749	A	G	1A, 1B
PS6	94258(Acc#AC004525)	35763	C	T	1A, 1B
PS7	94251(Acc#AC004525)	35770	G	A	1A
PS8	94204(Acc#AC004525)	35817	T	C	1A
PS9	94116(Acc#AC004525)	35905	C	T	1B
PS10	94077(Acc#AC004525)	35944	C	T	1A
PS11	94063(Acc#AC004525)	35958	G	A	1B
PS12	92691(Acc#AC004525)	37330	G	A	1A
PS13	92548(Acc#AC004525)	37473	C	T	1A
PS14	92435(Acc#AC004525)	37586	C	T	1A
PS15	92430(Acc#AC004525)	37591	G	A	1A
PS16	92417(Acc#AC004525)	37604	A	T	1A
PS17	92377(Acc#AC004525)	37644	C	A	1A
PS18	92343(Acc#AC004525)	37678	C	T	1A
PS19	86575(Acc#AC004525)	43446	G	A	1A
PS20	86318(Acc#AC004525)	43703	T	C	1A
PS21	77013(Acc#AC004525)	53008	A	C	1A
PS22	76922(Acc#AC004525)	53099	C	T	1B
PS23	76868(Acc#AC004525)	53153	T	C	1A
PS24 <sup>R</sup>	76608(Acc#AC004525)	53413	A	C	1A
PS25	76565(Acc#AC004525)	53456	G	T	1A, 1B
PS26 <sup>R</sup>	76516(Acc#AC004525)	53505	T	C	1A, 1B
PS27	76514(Acc#AC004525)	53507	C	T	1A
PS28	76508(Acc#AC004525)	53513	T	C	1A
PS29 <sup>R</sup>	76300(Acc#AC004525)	53721	T	C	1A, 1B
PS30	76106(Acc#AC004525)	53915	C	T	1A
PS31 <sup>R</sup>	76080(Acc#AC004525)	53941	A	G	1A, 1B
PS32	76072(Acc#AC004525)	53949	G	A	1A
PS33	75784(Acc#AC004525)	54237	C	T	1A
PS34	75553(Acc#AC004525)	54468	T	G	1A, 1B
PS35	75410(Acc#AC004525)	54611	T	C	1A
PS36	75323(Acc#AC004525)	54698	T	C	1A, 1B
PS37	75321(Acc#AC004525)	54700	T	C	1A
PS38	75280(Acc#AC004525)	54741	C	T	1A
PS39	75241(Acc#AC004525)	54780	C	G	1A
PS40	74938(Acc#AC004525)	55083	A	G	1A
PS41	74879(Acc#AC004525)	55142	G	A	1A
PS42 <sup>R</sup>	74693(Acc#AC004525)	55328	G	A	1A
PS43 <sup>R</sup>	74591(Acc#AC004525)	55430	C	T	1A
PS44	74482(Acc#AC004525)	55539	C	T	1A
PS45	74263(Acc#AC004525)	55758	G	A	1A

<sup>R</sup>Previously reported in the literature



## Example 2

This example illustrates analysis of the IL4R $\alpha$  polymorphisms identified in the Index Repositories for human genotypes and haplotypes for all polymorphic sites except PS1, PS9, PS11, PS21, PS22, and PS23.

- 5 A sampling of different genotypes containing these polymorphisms that were observed in these reference populations are shown in Table 4 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 4, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 4 can typically be inferred based  
10 on linkage disequilibrium and/or Mendelian inheritance.

Table 4. Genotypes and HAP Pairs Observed for the IL4R $\alpha$  Gen[illegible]

The haplotype pairs shown in Table 4 were estimated from the unphased genotypes using an extension of Clark's algorithm (Clark, A.G. (1990) *Mol Bio Evol* 7, 111-122), as described in U.S. Provisional Patent Application filed April 19, 2000 and entitled "A Method and System for Determining Haplotypes from a Collection of Polymorphisms". In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repositories examined herein and, by extension, the general population contains the 53 human IL4R $\alpha$  haplotypes shown in Table 5 below.

Table 5A. Haplotypes Observed for the IL4Ra Gene

HAP ID	PS 2	PS 3	PS 4	PS 5	PS 6	PS 7	PS 8	PS 10	PS 11	PS 12	PS 13	PS 14	PS 15	PS 16	PS 17	PS 18	PS 19	PS 20	PS 24	PS 25	PS 26	PS 27	PS 28	PS 29	PS 30	PS 31	PS 32	PS 33	PS 34	PS 35	PS 36	PS 37	PS 38	PS 39	PS 40	PS 41	PS 42	PS 43	PS 44	PS 45			
1	C	G	G	A	C	G	T	C	G	C	C	G	A	C	T	G	G	C	A	G	T	C	T	T	C	A	G	C	T	T	T	T	C	C	G	G	A	C	C	A			
2	C	G	G	A	C	G	T	C	G	C	C	G	A	C	C	G	G	T	A	G	T	C	T	T	C	A	G	C	T	T	T	T	C	C	A	G	C	C	G	A			
3	C	G	G	A	C	G	T	C	G	C	C	G	A	C	C	G	G	T	A	G	T	C	T	T	C	A	G	C	T	T	T	T	C	C	G	G	A	C	C	A			
4	C	G	G	G	C	G	T	T	G	C	C	G	A	C	T	G	C	C	T	C	T	C	T	T	C	C	G	C	T	T	T	T	C	C	A	G	C	C	G	A			
5	T	G	G	G	C	G	T	T	G	C	C	G	A	C	T	G	C	C	T	C	T	C	C	T	T	C	C	G	C	T	T	T	T	C	C	G	A	C	C	G			
6	C	G	G	G	C	G	T	T	G	C	C	G	A	C	T	G	C	C	T	A	G	T	C	T	T	C	A	G	C	T	T	T	T	C	C	G	G	A	C	C	A		
7	C	G	G	G	C	G	T	T	G	C	C	G	A	C	T	G	C	C	T	C	T	C	C	T	T	C	C	G	C	T	T	T	T	C	C	A	G	C	C	G			
8	C	G	G	G	C	G	T	T	G	C	C	G	A	C	T	G	C	C	T	C	T	C	C	T	T	C	G	A	C	G	T	T	T	C	C	G	G	A	T	C	A		
9	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	C	T	C	C	C	C	G	G	C	T	T	T	T	T	C	C	G	A	C	C	A		
10	C	G	G	G	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	C	T	C	C	C	C	G	G	C	T	T	T	T	T	T	T	T	T	T	T	T		
11	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	A	G	T	C	T	T	C	G	G	C	G	T	T	T	T	T	C	C	G	A	C	C	A		
12	C	G	G	G	C	G	T	T	C	G	C	C	G	A	C	C	G	C	A	G	T	C	T	T	C	G	G	C	T	T	T	T	T	T	C	C	A	G	C	C	G		
13	C	G	G	G	C	G	T	T	G	C	C	G	A	C	C	T	G	C	C	T	C	T	C	C	T	C	C	G	C	T	T	T	T	T	C	C	G	A	T	C	A		
14	C	G	C	G	C	G	T	T	G	C	C	G	A	C	C	C	C	C	C	T	C	T	C	C	T	C	G	C	G	T	T	T	T	T	C	C	A	G	A	T	C	A	
15	T	G	G	G	C	G	T	T	C	A	T	C	G	A	C	C	G	C	A	G	T	C	T	C	T	T	C	A	G	C	T	T	T	T	C	C	A	G	C	C	G	A	
16	C	G	G	G	C	G	T	T	G	C	C	G	A	C	C	T	G	C	A	G	T	C	T	C	T	T	C	A	G	C	T	T	T	T	C	C	A	G	C	C	G	A	
17	C	G	G	A	C	G	T	T	C	G	C	C	A	C	C	C	G	C	C	T	C	C	C	C	C	C	G	G	C	T	T	T	T	T	C	C	A	G	C	C	G		
18	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	A	G	T	C	T	T	C	C	A	G	C	T	T	T	T	T	T	C	C	G	A	C	C	A	
19	C	G	G	G	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	T	C	C	C	C	C	C	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
20	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	T	T	T	C	C	C	G	C	T	T	T	T	T	T	T	C	C	G	A	T	C	A	
21	C	G	G	G	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	T	C	T	C	T	C	G	G	C	T	T	T	T	T	T	C	C	A	G	C	C	G	
22	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	T	A	G	T	C	T	T	C	A	G	C	T	T	T	T	T	T	C	C	A	G	C	C	G	
23	C	G	G	G	C	G	T	T	G	C	C	C	G	A	C	C	T	G	C	A	G	T	C	T	T	C	A	G	C	T	T	T	T	T	C	C	G	G	C	C	G	A	
24	C	G	G	G	C	G	T	T	C	G	C	C	G	A	C	C	G	C	A	G	T	C	T	T	C	C	C	A	G	C	T	T	T	T	T	C	C	G	A	C	C	A	
25	T	G	G	G	C	G	T	T	C	A	T	C	G	A	C	C	G	C	C	T	A	G	T	C	T	T	C	A	G	C	T	T	T	T	T	C	C	G	A	T	C	A	
26	C	G	G	G	C	G	T	T	G	C	C	C	G	A	C	C	T	G	C	C	T	T	C	T	C	C	G	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
27	C	G	G	G	C	G	T	T	G	C	C	C	G	A	C	C	T	G	C	A	G	T	C	T	T	C	A	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
28	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	A	G	T	C	T	C	T	C	C	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
29	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	C	T	C	C	C	C	G	G	C	T	T	T	T	T	C	C	G	A	T	C	A		
30	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	T	A	G	T	C	T	T	C	C	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
31	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	A	G	T	C	T	C	C	G	C	T	T	T	T	T	T	C	C	G	A	T	C	A	
32	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	G	C	C	T	C	T	C	T	C	C	C	G	A	C	T	T	T	T	T	T	C	C	G	A	T	C	A
33	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	T	G	C	C	T	C	T	C	C	C	C	G	C	T	T	T	T	T	T	T	C	C	G	A	T	C	A
34	C	G	G	A	C	G	T	T	C	G	C	C	G	A	C	C	T	A	G	T	C	T	C	T	T	T	A	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T



In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

5 As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes  
10 prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:
  - (a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1, and the polymorphic variant comprises at least one polymorphism selected from the group consisting of guanine at PS1, thymine at PS2, thymine at PS3, cytosine at PS4, thymine at PS6, adenine at PS7, cytosine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, adenine at PS12, thymine at PS13, thymine at PS14, adenine at PS15, thymine at PS16, adenine at PS17, thymine at PS18, adenine at PS19, cytosine at PS20, cytosine at PS21, thymine at PS22, cytosine at PS23, thymine at PS25, thymine at PS27, cytosine at PS28, thymine at PS30, adenine at PS32, thymine at PS33, guanine at PS34, cytosine at PS35, cytosine at PS36, cytosine at PS37, thymine at PS38, guanine at PS39, guanine at PS40, adenine at PS41, thymine at PS44, and adenine at PS45; and
  - (b) a second nucleotide sequence which is complementary to the first nucleotide sequence.
2. The isolated polynucleotide of claim 1 which comprises an IL4R $\alpha$  isogene.
3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.
4. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an IL4R $\alpha$  protein encoded by the first nucleotide sequence.
5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.
6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the IL4R $\alpha$  gene, the fragment comprising one or more polymorphisms selected from the group consisting of guanine at PS1, thymine at PS2, thymine at PS3, cytosine at PS4, thymine at PS6, adenine at PS7, cytosine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, adenine at PS12, thymine at PS13, thymine at PS14, adenine at PS15, thymine at PS16, adenine at PS17, thymine at PS18, adenine at PS19, cytosine at PS20, cytosine at PS21, thymine at PS22, cytosine at PS23, thymine at PS25, thymine at PS27, cytosine at PS28, thymine at PS30, adenine at PS32, thymine at PS33, guanine at PS34, cytosine at PS35, cytosine at PS36, cytosine at PS37, thymine at PS38, guanine at PS39, guanine at PS40, adenine at PS41, thymine at PS44, and adenine at PS45.
7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IL4R $\alpha$  cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 237, adenine at a position corresponding to nucleotide 244, cytosine at a position corresponding to nucleotide 291, thymine at a position corresponding to nucleotide 501, adenine at a position corresponding to

nucleotide 554, cytosine at a position corresponding to nucleotide 939, thymine at a position corresponding to nucleotide 1242, thymine at a position corresponding to nucleotide 1293, cytosine at a position corresponding to nucleotide 1299, thymine at a position corresponding to nucleotide 1701, adenine at a position corresponding to nucleotide 1735, thymine at a position corresponding to nucleotide 2023, guanine at a position corresponding to nucleotide 2254 and cytosine at a position corresponding to nucleotide 2397.

8. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses a Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) protein encoded by the polymorphic variant sequence.
9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.
10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the IL4R $\alpha$  protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO: 3 and the polymorphic variant comprises one or more variant amino acids selected from the group consisting of threonine at a position corresponding to amino acid 82, histidine at a position corresponding to amino acid 185, isoleucine at a position corresponding to amino acid 579, serine at a position corresponding to amino acid 675, and alanine at a position corresponding to amino acid 752.
11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.
12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the IL4R $\alpha$  polymorphic variant with a candidate agent and assaying for binding activity.
13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene at a polymorphic site selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45.
14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IL4R $\alpha$  gene at a region containing the polymorphic site.
15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-79, the complements of SEQ ID NOS: 4-79, and SEQ ID NOS:80-231.
16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.
17. A method for genotyping the Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene of an individual, comprising determining for the two copies of the IL4R $\alpha$  gene present in the individual the identity of



the nucleotide pair at one or more polymorphic sites (PS) selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45.

18. The method of claim 17, wherein the determining step comprises:

- (a) isolating from the individual a nucleic acid mixture comprising both copies of the IL4R $\alpha$  gene, or a fragment thereof, that are present in the individual;
- (b) amplifying from the nucleic acid mixture a target region containing at least one of the polymorphic sites;
- (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
- (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and
- (e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

19. A method for haplotyping the Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene of an individual which comprises determining, for one copy of the IL4R $\alpha$  gene present in the individual, the identity of the nucleotide at one or more polymorphic sites (PS) selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45.

20. The method of claim 19, wherein the determining step comprises:

- (a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the IL4R $\alpha$  gene, or a fragment thereof, that is present in the individual;
- (b) amplifying from the nucleic acid molecule a target region containing at least one of the polymorphic sites;
- (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
- (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and
- (e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene of an individual comprising:

- (a) identifying an IL4R $\alpha$  genotype for the individual at two or more of polymorphic sites selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45;
- (b) enumerating all possible haplotype pairs which are consistent with the genotype;
- (c) accessing data containing the IL4R $\alpha$  haplotype pairs determined in a reference population; and
- (d) assigning a haplotype pair to the individual that is consistent with the data.
22. A method for identifying an association between a trait and at least one genotype or haplotype of the Interleukin 4 Receptor Alphagene which comprises comparing the frequency of the genotype or haplotype in a population exhibiting the trait with the frequency of the genotype or haplotype in a reference population, wherein the genotype or haplotype comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS1, PS2, PS3, PS4, PS6, PS7, PS8, PS9, PS10, PS11, PS12, PS13, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS21, PS22, PS23, PS25, PS27, PS28, PS30, PS32, PS33, PS34, PS35, PS36, PS37, PS38, PS39, PS40, PS41, PS44, and PS45, wherein a higher frequency of the genotype or haplotype in the trait population than in the reference population indicates the trait is associated with the genotype or haplotype.
23. The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-53 shown in Table 5.
24. The method of claim 23, wherein the trait is a clinical response to a drug targeting IL4R $\alpha$ .
25. A computer system for storing and analyzing polymorphism data for the Interleukin 4 Receptor Alphagene, comprising:
- (a) a central processing unit (CPU);
- (b) a communication interface;
- (c) a display device;
- (d) an input device; and
- (e) a database containing the polymorphism data;
- wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4 and the haplotypes shown in Table 5.
26. A genome anthology for the Interleukin 4 Receptor Alpha(IL4R $\alpha$ ) gene which comprises IL4R $\alpha$  isogenes defined by haplotypes 1-53 shown in Table 5.

POLYMORPHISMS IN THE IL4R $\alpha$  GENE

TGTGAGCTAC	TGTGTCTGGC	CTGAATAATA	AAATTTAAAA	CAATTTTTC	
AAAATTCACC	ATGAGGTCTC	ACTATATTCC	CTAGGCTGGT	CTCAAACCCC	30100
TGGACTCCAA	GTGATCCAAC	CCACCTTCCC	GAGTAGCTGG	GACTAGAGAT	
GCACACCATT	GCACCCAATA	GAGCAATACG	TTTCTGTTCT	TTGTAAATTA	30200
CCTGCTCTAA	GGTATTTTTG	TTATAGCAGC	CTATATGGAC	TAAGCTGACT	
TGTAACGTTA	CTTGAGACTT	TAAAGTGTTT	CGGTCACTGT	TGGAGGGCTC	30300
TGTCTGTGTT	AGCTCATTTA	ATCCCCACAA	CACCTCAATC	AGATGGGGCT	
ATTCTTAGTC	CCACTTTATA	GATAAGGAAA	CTGAGGCATG	GAAGCACAGC	30400
TTGCTCAAGG	TTCACATCTA	GTCAGTGACA	GAGCAGGTAT	TTAAACCTCA	
GGAAATAATC	AGAGAAACAT	GTGTAGAGGG	TTGTCCAAGG	AAGGCCACAT	30500
CCAGAAGCAT	CTCCCAGGAC	AGTTGTTGTG	TAGCTCACCC	TCTGGACTTT	
GTGGGTCTGG	GTGTTGTTTC	ATGATTATAG	AGAGAGCTCT	GTGAACGTGG	30600
AGGACCTGTT	GTCGGCAGAG	ACACAAATGG	CCAGGGCATG	GCTGGGCAGC	
CGCAGTGGCT	CAGGCCTGTA	ATCCCAGCAC	TTGAGAGAAG	CCAGAGGGGC	30700
AGATCATGAG	GTCAGAAAGT	CAAGACCAGC	CTGGCCAACA	TGGTGAAACC	
CCGTCTCTAC	TAAAAATACA	AAAATTAGCC	AGGTGTGGTG	GTGGGCACCT	30800
GTAATCCCAG	CTACTCGGGA	GGCTGAGGCA	GAAGAATCGC	TTGAACCCGG	
GAGGTGGAGG	TTGCAGTGAG	CTGAGATTGC	ACCCTGCAC	TCCAGCCTTG	30900
		G			
GAGACAGAGC	GAGACTCTGT	CTCGGAAAAA	CAAACAAACA	AGCAAACAAA	
CAAACAAATA	AATGGCCAGG	GCAGGGGAGG	GTTGCATATT	GAATAAGATG	31000
AGCTCTGCTG	GAAGCACAGG	TCAGCACTAA	CCTGCTTCCT	CTCTCTCTGC	
AGGTGCCTTG	GCATCTCCA	ATGGGGTGGC	TTTGCTCTGG	GCTCCTGTTT	31100
	[exon 3: 31071..				
CCTGTGAGCT	GCCTGGTCCT	GCTGCAGGTG	GCAAGCTCTG	GTAAGTCACC	
	..31140]				
ACTTCTCAAT	CATTCAATTG	TTGGCTATTA	ATGGCGTGCC	AGGGTCCTGC	31200
AGTATGTCAC	CTGGCCTTAT	GGAGATTACA	CTGCAGTGGG	AGGGGACAGC	
CAATGACAAG	TGGCCCTGAT	TATCAGTAAA	TTCTAAAGAT	TGTTAGAAAG	31300
TGATGGGAGC	CGGGTGCACT	GGCTCACACC	TGTAATCCCA	GCACTTCAGG	
AGGCCGAGGC	AGGAGGATCG	CTTGAGCCCA	GGAGTTCGAG	GTCAGCTTGG	31400
GCAACATAGG	GAGACCTTGT	CTCTACAAAT	AATAAAATAT	TAGCCAGGTG	
TGGCAGTGCA	CGCCTGTAGC	CCCAGCTACT	CAGGAGGCCG	AGGTGGGAGG	31500
ATCCCTTGAA	CTCAGGAGGT	CAAGGCTGCA	GTGAACTGTG	ATCGCGCCAC	
TCCACTCCAG	CCTGCGTGAG	AAAGTGAGAC	CCTGTCAAAA	AAAAAGAGAA	31600
GGTGATGGGG	AAAGAACACA	GAACAGCATA	AGAGGGGGTT	GGGGAAGCTG	
GGTGGAGTGG	GGGGGATTGC	AGTTGAAAGT	AGGGAAGTCA	GGGAAGGCCT	31700
CATTGAGCTG	ACTTGGAGGA	AGCGGGAACC	GTGCAGATGT	CTGGGGAAGG	
CTCATTCTTG	GCAGAGAGGC	CCTGCACTGA	GCCTGGCGGG	AGGGTTGAGC	31800
ACAGGAGGGA	ATGTGGTGGA	GGAGAGTGAG	CAGCAGGAGG	GAGCAGTGAA	
GGTCAGCAAG	GTGACAGAGT	GGCTGAATCA	AAAAAGACCT	TGCAGTGTTT	31900
GAGCAGAGGA	TCCATATCAT	CCATTATGTT	CCAAAGGACT	CTTCAGGATG	
CCGTGTGGAG	AAAGGAAGAG	GGTGAAGCC	AGGAGGTCTG	GAGGGAGGTC	32000
TGGAGTGGAG	GAGATGAGAG	GCTCCGGATC	CCTCTGGGAG	GTAGATTTGA	
GGACAGATTG	GAATTGAGGT	GAAAGACAGA	GAAAGAGAAG	TGGCCAGGAT	32100
GACTCCAAGA	TTTCTGACCT	AAACTACTGG	GAAGGACGCG	GTGTGCATTT	
CTGAAATGCA	GAAGGATGCC	AGAAGAGAAG	GTAATTTGGG	GAGGGGCGGG	32200
AATCAGGAGT	TAGTTTTGGA	CATGAGATAA	GCTTGAATA	TTTATTTGCT	
ATCTAAGACA	GCTCCTTAAC	ATGGTAAGCC	CTTATGCAAG	TTGTTGTCAG	32300
CTGAGATGGG	CGTGGCACTG	AGCATGGGAG	CATGGAGGCG	CCTGAGTGST	
CTCATGCTCA	GGTGGTTTAG	CAAACTCAGT	GTACATCCTG	CCAATTCCAG	32400
TCTTGCCATG	GCCACTGACA	AGCTAGGAGG	GCGCTGAAAG	GAGAAGGACC	
CCGATGTCTC	CTCCAGCCCA	TCCATCTCCT	CTCTCCCAT	GGCCAAACCC	32500
AACCGGAAAC	TAAAGGCCAA	GGGTACCCGG	TGATGAAGAC	TGTGGTATCA	
GCCTCCTGAG	CACAGAGAGG	GCAGAAAGGG	GTGGAGACAA	AGAGGGGCGC	32600

FIGURE 1A

AGATAGTGGG	CAAATGGGGA	AGTGGCACTT	CCCCTAGCTC	GAGGGCAGAG	
GCTTGGTGTG	ATGGAATGGC	ACTCCTTAAA	CTGCTACATA	TTTTCCCTTT	32700
AATTTGGCCA	AGAACAAGTT	GTCAAGTTTG	TGTGAGATAA	AGGTGCACTT	
GGTTCGTTCT	TGTCTAATGG	CCCCCGCACC	CATGGGTATT	TCTTCAGCTT	32800
CCACAGTCAT	CCCACACTA	GCTGGGAAGC	TCCAGCAGCC	CTGGTCCTGG	
CCCCAGCTCT	GTGGGCGCTG	GCCCTCAACT	TTGCCTGCAC	TGTGCTTTTG	32900
TGCTATTCCC	CTTGGTCCTG	TTTGGGTGCA	AGTCCCCCTC	ACGCATTGAG	
T					
TTCTTGGGCC	GCTCAGGCTG	CTCCTGTGTC	TCCCCAGGGA	ACATGAAGGT	33000
	T [exon 4: 32988..				
CTTGCAGGAG	CCCACCTGCG	TCTCCGACTA	CATGAGCATC	TCTACTTGCG	
AGTGGAAAGAT	GAATGGTCCC	ACCAATTGCA	GCACCGAGCT	CCGCCTGTTG	33100
TACCAGCTGG	TTTTTCTGCT	CTCCGAGTAA	GCCTGCGCTG	GAGCTGGAGG	
	..33126]		C		
TTTGGGGAGG	TTGTGCCCAA	AGGGTTTGCC	CCAAGAGTGA	GCTGGGTCCA	33200
GGTGGTGC GC	TGGAGTG CAG	GATGCTGAGT	ATGGTTTGCT	GCTGTTTATA	
TGGTGTTAGA	GGGGAGGTCC	CATCTCCAGG	GACATGTTAT	GTAAGATACA	33300
GTGGAGCGCA	TGGTGGGAGT	GTTGGTCCAC	GTGGCACATG	GATACGGCTG	
GAATACTGGA	CTAGACCAGC	AGTTCTCACA	CTTTTTGGTC	TCAGGACCCT	33400
TTTTTCACACT	TAAAAATGAG	TGAGGACCCA	AAGGGCTTTG	GTGTAGGTAA	
CACATCATTC	TATGTTTACC	TAATTAGAAC	TTGCAATGAA	GAAATGGTGT	33500
AATTTTAAA	AAATTA AAC	AATTA AAAAT	TTTTTTTCTT	ACTGAAATGG	
AGGTCCTACT	GTGTGCCCCA	GGCTGCTCTC	AAACTCCTGG	GCTCCAGTGA	33600
TCCCTCCTGCC	TCCGCCTCCC	AAAGTGCTGG	GATTACAAGC	GTGAGCCGCT	
GTATCCGGCC	CAAATGGGAG	AAATTTTAAG	TCCCAACAAC	ATGCAAGCCC	33700
GCATTCAACA	AATCTTCAGA	TCAATTACAT	GATCACAGGT	CATGTAGCCT	
CTAGAAAATT	CCACTGTACG	CCAGTGAGAG	AGAGTGAAAA	GGCAAATAAC	33800
GTCCCCTGTAT	TATGATGAAA	AGAGTTTTAC	CTGGTGGGCC	CAGACCACAC	
TTTGAGAACC	ACTGGACTAG	ACCCTTGATT	GAGGAGTACG	GTGTTGAGAG	33900
TGGAGTCCTC	TGTGATGGTG	GATGGACCAG	GACACATGGC	ATAGGAGTCA	
GGTGGTTCCC	TGGGCTACTC	CATGGTGCAC	AGGATGCTTC	GTTACACTGG	34000
TGCCCAGGAC	ATAATCAGT	ACACAAGACA	CACAGTTACG	GGGCAGACTG	
GGGATATACG	GCACACCAGC	ATGCAGCGTT	CACCAGTAAA	GGTGGTATTC	34100
CATGATTATT	CTAAGGTAGA	TGGGCTGTGC	TTTGTTTCCA	TTGGCTTAGT	
CCAGGGATTG	GCAAACATATG	GCCCCGTGAGC	CAAATCCGGC	CCACTGCTTG	34200
TTTTTGTA AA	TAAAGTTT TA	TTGGAACACA	CTGGCTGCTG	TAGTTGTAAC	
AGAAACTGCA	TGGCCCTCCT	TTATGTTTTT	TGTTTTGTTTG	TTTGTGTTGTT	34300
TGTTTTCTTT	GAGACAGAGT	TTCGCTCTTG	TTGCCCAGGC	TGGAGTGCAG	
TGGCACAA TC	TCGGCTCACT	GCAACCTCTG	CCTCCCGGGT	TCAAGCGATT	34400
CTCCTGTCTC	AGCCTCCCGA	GTAGTTGGGA	TTAATGGTGC	CTGCCACCAC	
ACCCGGCTAA	TTTTTCGTAT	TTTTTAGTAGA	GACCGGTTTT	CATCATGTTG	34500
GCCAAGCTGG	TCTCGAACTC	CTGAACCTCAG	GTGATCCACC	CGCCTCAGCG	
TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	CACTGAGCCC	GGCCTCCTCC	34600
TTTATCTTAA	TTGAAATAAT	TCAGAAATGG	AAAGTCAAAT	ACTGCATGTT	
CTCACTTATA	AGTAAGAGTT	AAATAATGTG	TACACATGGG	CATTATTCCA	34700
TGTACCATGG	AATAACAGAC	ATTGAAGACT	TGGGAGGGTG	GGAGAGGGGT	
GAAGGAAGAG	AAGTTACTTA	ATGGGCATAG	TGTACACCAT	TTGGGTGACG	34800
GACCCACCAG	AACCCAGAC	TTCACCACTA	GGCAGCATAT	CCAGTGAGAA	
CAGATCTGAG	GCTTGCCATC	AAAATTGCAC	TTGTAAGGCC	GGGCACTGTG	34900
GTGGCTCGCG	GCTGTAATCC	CAGCCCTTTG	GGAGGCCGAG	GTGGGCAGAT	
CACTTGAGGT	CAGGAGTTCCG	AGACCGGCCT	GGCCAACATG	GTGAAGCTCC	35000
ATCTCTACTA	AAAATACAAC	AATTA ACTGG	GTGTAGTGGC	GCACACCTGT	
AATCCCAGCT	ACTAGGGAGG	CTGAGGCGGG	AGAATTGCTT	GAGCCCAGGA	35100
GGTGGAGGTT	GCAGTGAGCC	GAGATCACAT	CACTGTACTC	TAGCCTGGGT	
GACAGTGAGA	CTTTGTCTCA	GGAAAAA AAA	ACAAAAACAA	AAAAACAAAA	35200
ACTCGTACCC	CCTAAATTTA	TACAAATAAC	CAAAAAA AAA	AAAAAAAAG	
GAAATTGTGT	GGCCTTTGAA	GTCCAAAATA	TTA ACTATCT	GGCCTGTTAC	35300

FIGURE 1B

AGAAAAAGTT	TGCAGACCCC	TGGCCTAGCC	CGTGAGATGT	GGGTTGGCTG	
TTAAGGTGGA	ACATTGGAAT	TATCTTACGA	TGGCCAAACT	GTGCGATGCA	35400
GAGCTTATGT	TGTTCTAAAT	TAATTAGTGC	CACCGGTTCT	TCCCTTTCAT	
GGGCTTTTCA	GAACAAGCTA	AGTCCCAGGA	CCAGGGCCGG	CAGCTAGGCA	35500
GGTGTGAGGA	GCATCCTTGG	TGCATGTGGT	AAGAGGCTGT	GGCCAGCAAG	
AGAGGCAACC	CTAGTCGGCT	GCCCCAGCAC	ACCCTGGCCG	CTCCCAAGCC	35600
CCCAGATCTG	TCCTCACATC	CGTGATCGGG	AAGCTGGAAG	AGTCTGATGC	
GGTTCCTGGA	GGCATGTCCC	GGACACAGCT	GTGGGGCCCA	GCCAGCCTAC	35700
AGGTGACCAG	CCTAACCCAG	CCCCTGTGTC	TGCAGAGCCC	ACACGTGTAT	
			G		
[exon 5: 35736..					
CCCTGAGAAC	AACGGAGGCG	CGGGGTGCGT	GTGCCACCTG	CTCATGGATG	35800
	T	A			
ACGTGGTCAG	TGCGGATAAC	TATACACTGG	ACCTGTGGGC	TGGGCAGCAG	
	C				
CTGCTGTGGA	AGGGCTCCTT	CAAGCCCAGC	GAGCATGGTG	AGCAGGGCGG	35900
	..35887]				
AGTGCGGCAG	GGGTGGCTGG	GTGTGTTCCC	ACAGCTGCCT	GGGCTGAGGG	
T				T	
TGGGGTGGGC	AGGGGAGGAG	GTGGGGTCAT	AGCAACAGCA	GGAGGAAGCC	36000
A					
GCCTGTATTT	TCCCAAATCT	GATGGGATTC	CTGCCCCGTC	CTGGGGCTCA	
GTCCTCCAC	CTTTGAAACG	GAGCTGGTCG	CAGTAGACCA	CCAAGCCCCC	36100
TTCAGCCCAG	CTGTTTCCAC	CCCTGAACCT	AAGTGCCAG	GAAGGCGTAT	
TGAGATGAGG	TGTGCTTGCT	GGAAGGCATG	CCTGCTGCTG	ATTGAAAACC	36200
GAAGTGGGAA	CATTCCCTCC	ATTCTGTGTC	CACTGGTCAG	CTGCTGCGGC	
TTTGGATGGT	CTTGACCGTG	GAAGGCTGAC	CTTCTTCTGG	TACCCGGAGT	36300
CCCTGCAGGA	ATCCCCCTTG	AGCTTGCTGG	GCTGTGGTGA	CAGGAGTTTA	
AAACATGCGT	TGTATTCCAG	TGATGCATGA	TATGACATGC	ATCACAGGAA	36400
TAAAAACCTG	AGGTCTCATG	GATATGATTG	CTTCAAAGGA	GACCAAGTTT	
TAAACAGAT	GAATCAAAAT	AAAGAAAAAT	ACTCAGTAAA	TCATCATAAA	36500
GTACAGAGAT	GTGGCCAAAG	GTGTGAAGGA	TGCAGCTGTA	AAAGCTGAAG	
TTTGAGGCCG	GGTGTGGTGG	TTCATGCCTA	TAATCCCAGC	ACTTTGGGAG	36600
GCCGAGCCCA	GCGGATCACC	GGAGGTCAGG	AGTTCGAGAC	CAGCCTGGAC	
AACATGGTAA	AACCCCGTCT	CTACTAAAAA	TACAAAAAAT	TAGTCTGGCA	36700
TGGTGGCAGG	CGCCTGTAAT	CCCAGCTACT	TGGGAGGCTG	AGGTAGGAGA	
ATGGCTTGAA	CCCAGGAGAA	GGAGGTTGCA	GTGAGCTTAG	ATCATGCTAC	36800
TGCCCCCCAG	CCTGGGCGAC	AGAGTGAGAT	TACGTCCTCA	AAAAATAAAA	
ATAAATAAAA	ATAAAAAGAT	TTTTTAAAAA	GCTGAAGTTT	GGGTTACTTT	36900
GGCTCATACA	CTTTGCCTTC	ACTGTAGAAA	GGTGGTTAGT	AAAGACCAGG	
CGCGGTGGCT	CATGCCTGGA	ATCCCAGCAC	TTTGGGAGCC	CAGCGCAGGC	37000
AGATCACTTG	AGCCCTGGGC	TATTGAGGCT	GCAGTGAGCT	GGGATTGTGC	
CACTGCACTC	CAGCCTGGGC	AACAGAGTGG	GACCCTGTCT	CAAAAAAGAA	37100
GAAAAAAGG	GTAATTAATA	AACACTAAAG	TTCTATGTAG	AATTTTAGCA	
ACATTATTGT	TATTATAATC	TTCTTTGCTA	TGGCTCTGAA	TCTGTGTGGT	37200
GCTCCAGAAG	TATGCTATGG	AGGTTTTGTC	GACCAAAAAA	CTGGGTGGTG	
GCTGTGGTTT	GTAGGCCGGG	GCTGGGCTGG	GTGATGGGGG	AGTCACTGCA	37300
TAGATCCTCA	CATAGAGGCC	GCTTCTCCCC	CAGTGAAACC	CAGGGCCCCA	
		A			
[exon 6: 37334..					
GGAAACCTGA	CAGTTCAACAC	CAATGTCTCC	GACACTCTGC	TGCTGACCTG	37400
GAGCAACCCG	TATCCCCCTG	ACAATTACCT	GTATAATCAT	CTCACCTATG	
CAGTCAACAT	TTGGAGTGAA	AACGACCCGG	CAGATGTGAG	TGGGCATGCT	37500
		T			
	..37485]				
TTGACGTTTT	TCTGTGACCT	CTGGGGAAACA	GGGTGGGTGA	CCAGCAGAGG	

FIGURE 1C

CCCAGTCCCT	GGAGCCAGGA	GCCCTGGGAGG	CAAGCCCTGG	GGCTGGATAG	37600
CAAATCCCAG	GAGCTAGAGA	CCTGGCTTCT	T CACCTGGCTC	A TGCCCTAGGC	
T AAGTCCCTTT	GCTTCCTGGC	CCCCCACCCC	T TCACATCAGA	A GAAGGGGAGT	37700
TATCTCTGCA	TGCCGCTCCT	CCTCTGTAAA	GGTAGGGCTG	TGGGCCACAT	
CTGTGTTTCC	CAGTTTGGGG	GACACAAGTG	ATCGTAGGTG	GCACATTGAC	37800
AGCTCACTTG	AATAACCCTA	TTATTGAAGA	GAATAATACT	GACTCAAGAG	
ACAGTGACCC	GTGTCACTTC	CCTTTTGAGG	CCAACGGGTT	AAGGAGGAAG	37900
TCCCCATACA	GCTGACTCGT	TTACTAATTC	CTCTTAATGA	AGAGAGCAGA	
GGCCACACCC	CAGGCTTAGA	CTTTCCCAAG	AAAACAAGAT	CAGTTTGTTG	38000
GTTGTTCCCC	ATGGAAGCTG	GTCTTGACAT	TCCCTTCACA	GTAGTGTGG	
TGGAGTTT	GTTGTTGTTT	GTTTTGAGAC	AGAGTCTCAC	TCTGTCACCC	38100
AGGGTGGAAC	ACAGTGGCGT	GATCTTGGCT	CAC TGCAACC	TCCGCCTCCT	
GGGTTCTAGC	GATTCTCCTG	CCTCAGCCTC	CTGAGCAGCC	GGGACTACAG	38200
GCACCTGCCA	CCGTGCCCAG	CTAATTTTTG	TATATTTAGT	AGAGATGGGG	
TTTCACTGCG	TTGGCCAGGC	TGGTCTCAAA	CTCCTGACCT	CAGATGATCC	38300
ACTCGCCTTG	GCCTCCCAA	GTGCTGGGAT	TACAGGTGTG	AGCCACCGCA	
CCTGGCCAGT	GGAGTTCCTT	CTTAAGTACA	TGTATTGACA	TCTTTAAAAA	38400
GGGCGAGAGG	ATTTACAGGA	AACATACAGG	TCAGTAATGG	CAGGGGCCGT	
CCACAGTGGG	TGGCTGAGTC	CCCCATTTTT	TCTGCTGGTG	TGCAGGGAGG	38500
TCATTTCCCTG	CCACCCATGT	TTCCCCACCC	TGAATCCACC	TTCCCTCACAT	
TCCCATTGGA	GGGACAATCT	CTGGACATAT	GGGACCTGGG	GTCCACACAGG	38600
GCTGCAATCC	AATGCCTGCT	GTGCCACTCG	CCAGCTGTGT	GATGTTGGGC	
ATATCCCAT	ACCTCTTTGT	GCCTCAGTTT	CCTCATCTGT	AACACAGGAG	38700
TGACAAGAGC	ACCCGCCCAC	AGGGCTATGA	CAGTACAAGG	TGTGTGATAC	
AGATGAGCTC	CCCTGTTTGG	CCCACATGTG	TCCTAAAAGC	CATGTGCCCT	38800
TTCTCTTGAG	TGCCCCAGGC	CACAGAGATC	CCCATCTGCC	CGCTGTCCCA	
CACACTGGTC	TGTCATTTGT	TCCTTGAGGT	TTGTGAGGGC	CGGCTCTGTG	38900
CATCCCAGGG	GCCCAGGCTG	GGCCTGGTTG	GCTCTCAGGG	AGCAGGCACC	
CGCCACCTTA	AGCTCCCATG	CTGGTGTCTG	TCACTGCTTC	CTCTCAATCT	39000
GGCCAAGCCA	GGGGTGTCGA	TTTATATCTC	TCAGGTCTGG	TTTCCCCTTT	
GGCACTGGGC	CAGGTATGGG	GAAAGAGCAG	GAATGGGGCA	GTTGGCTCAC	39100
ACAGCAGAGG	CTCAGAAAGC	GGGGGGCATG	GGGGGAAGGA	GTGCACAGAT	
GCTAGAGAGT	GGGGCAAGTT	TTGTTTGGTC	AATAAATCTC	CTTCTCATGC	39200
CCCAGGCCCTG	TGCAAGACCT	ACAGAGAGTC	CCAAGGATGG	GCTGGGGGGA	
AGAGAAAGGT	ACCACCTTCA	GAGTCCAAAG	ATATGTTATT	TAATATTTTC	39300
ATATTTCTAG	ATCTGCCCTC	AGGCATGGCT	GGATCCAGCT	TCTAGGAACC	
TGTCCAGCTC	TGCGCCCTGC	TTTATCTGT	ATTGGCTTCG	TTTTTAGGCA	39400
GGCTCTTCCC	TCATGTAGTG	GCAGATATGC	CTACTAGTTG	CTCCAGGCCT	
ACATCCCAA	GCCACAGTGG	GAAAAGGGTT	TTTTTTCTTG	ACGGTTCCTAA	39500
TAAGAGTCCT	AAGGCTGCTG	CTCAGTGGCC	TGGCTTCGAT	GCTGTGCCAG	
CCTCTGAACC	AATCACTGGC	TGTGGGTGGA	GAGAGGGTGC	TGGTGGAGGG	39600
CCCTGCTTGT	CCAGGGAGGA	GTACATACC	TGCCTCTAGG	GCTGCAGGTG	
GGCTCAGCTC	CATCCAAACC	AGATGAACTG	AAAAATAAGGC	AGGAGTGGCT	39700
TCCCCAGGGG	AAACTGGGGA	AGAGGAAGCA	GGACTGTGCT	GGCTAAAATG	
CCAGCCAGGT	TTAAGACGTG	GCACCAGATG	CCAGTCATGG	GATTGGATTG	39800
GTCAGCATGC	CTGGGCTATG	GCTTAGGGGT	ATGTTGGTGC	TCAGGGATGC	
CACAGGCCCTC	CAGTACCAG	GTCTGAGGCA	GAAGAAATGAA	GTCCAGCTTC	39900
TCTTGTTGGGT	GGAACAGTGG	CAACTGAGAT	ACCCCATCTC	TCCCTTCCCA	
AGAACAGAGC	TGAACATAAA	GAATTTAGTG	ATTGGCCAGA	GCTTGGCCAC	40000
ATGCTCCCCCT	CTGATGAATG	ATAGGCCAGG	TGATGGGATT	GGCACAATTG	
GCTTAGACTA	ATGAGGGTTG	GCCCTGGAGT	TGCAGGCAGT	GGAGTTCTGT	40100
CCTAAGCAGT	GGGCACCTAA	ACCCGATGGC	ATAAAAGCTG	GGCGGGTGTC	
CACCTGCATC	TGCCACAGCA	CTATAGGCAC	CAACTGTGGC	TCATACTGAG	40200
TGGGATAAAT	TCCAGAAAGA	AACATTAGGA	ACTTACTATA	GAATTTTGGG	

FIGURE 1D

GCTAGAGCTA	CTCATTTCATT	CCCCTAGATA	ATTTCTAGGC	AAGGTTCCAT	40300
AGTGGAGGGG	GAGTTTTTGGC	TTGGGGCATTG	AAGGATGCAT	AGGAGTTTTTC	
TAGATGGGGA	AAGAAGGGAA	CGGTAGACCA	GGCAGAGGGA	ACTGCATGAT	40400
AAAAGGTTTA	TGGGTGTGAA	AATTCATGGA	ATGTTTGAGG	ATTATGGGGT	
TGGGGGATGT	GGGAATATGT	GTAGCGATAA	AGCACCAAAC	AAAGCCAAAA	40500
GTTTAGTTAG	AGCCTGAAT	GCCTGCCCTGA	TAATGGTTTC	CATATTTTAT	
ATGCCCTACTA	TGTGCCAGGC	ACATTGCTCA	GGGTCACACA	GCTGGAAATG	40600
GCAGGGCTGA	GTTTTTGTGT	TTGTTGTTGT	TGTTGAGACA	GAGTCTCACT	
CTATCACCCA	GGCTGGAATG	CAGGGGCGTG	ATCATGGCTC	ACTGCATCCT	40700
TGACTTCCTG	GGATCAGGTG	ATTCTCCCAC	CTCTGCCTCC	CAGGTAGCTG	
GGACTACAGG	CACAGGCCAC	CACGCCAGGC	TAATTTTTTG	TATTTTTTAGT	40800
AGCGACAGGG	TCTCGCCATG	TTGTCCGGGC	TGGTCTGGAT	CTCCTGGCTT	
CAAGTGATCC	CCCTGGCTCA	GCCTCCCAAG	GTGCTGGGAT	TACAGGCTTG	40900
AGCCACCGCA	TCCAGCCCAG	ATCTGAGATT	TGCACCCAGT	ATTTGAACCTC	
CCAAGCCTGT	GCTCTTTTTC	CTCCCATGGA	CATTTCTCTC	AGAGATGGTC	41000
TCCCAAACAC	CTGTCCTTCT	TGTTAAAAAA	CAGACAAACC	GCAAGTAGTT	
CTTTGGAAGC	TCAGATTTCT	CTTTTGTTC	TTAGTAAAC	ATTTCCAGT	41100
TCCCAGCTCC	CTTCCAGGTG	GTAAGATTTT	TTCCGTAAC	TACATCTAGC	
TGTTGCTTCT	TGTTTGCTCA	TGTTTAGAAA	GAAAGACAAA	AGAGAGTGAG	41200
AATTTTCTCT	CCCTTCCCCA	GTCTCCCCAC	AACTCACACC	CCACCCTCAG	
CTCCCTCTGT	AATAGGAAAA	TCTCTGAACT	CTCTGTAGTT	GCTCCAGCAA	41300
TCTTTTGGA	CTTTGCTTCT	TTCTTGTA	AAAACCTCCC	CTTGGCTCAC	
TTTGCAACAG	GTTTCCCCAA	ATGTGCTTCC	AACCACAAGC	AGAAATGGAG	41400
CTGCCAGTAA	CAGGGAAGAA	ACTGCCGGGG	GCTGAGGAAG	AGGAGAGGGA	
GGTGATAGC	CCTGGATCTC	GCAGGGAGAG	GGGTGACAGG	ATGAGAAGTC	41500
AGGTTGCTCA	CTTGCCATCA	GGGTGAGTCA	TGAATATAGC	GTTTCATGTAT	
CACTTTTTTAA	AGCTTTTTTG	GAGGGTAAAA	GTAATAGTTA	CACAAAATAA	41600
AAATACAAAT	GGTACAAAAG	GACTTAGAAT	GGAAACATGT	TTCTCTCCCG	
ACTCCAGCCT	CCTGTTTTTC	TTCCAGAGA	CTGACCACTG	CTGTCTGTCT	41700
CTTGCCAGAA	GGGAAAGGGA	GGCAAGGTTA	GGGCAGGCAG	AGGGCATGTG	
CATCCTTTAG	AGAGAGCTTA	TGTCTATACA	AGCAAATGTG	TGTGTTTCAGT	41800
CATCGTGTG	TTAGTTTTCT	ATTGCTGCAT	AATAATGGTA	CTACCAGCTT	
CACAGCTTTA	AACAACACCC	ATTTATTATC	TCATAGTTTC	TGTGGTTGGG	41900
AGTCTGGACA	TAGCTTAGCC	AGGTTCTCTG	CTTTAGAGTC	TCGTGAGGCT	
ATAATCAAGG	TGTGGGATGG	GGCTGCAGTT	TCATCTGAGG	CTCAATTGGG	42000
GAAGGGTCAC	TTCTAAGCTC	ATACAATATT	GGTGACATTC	AGTCCCTGGC	
AGGCTGTTGA	ACTGAGAGCC	TCAGTTTCGT	GCTGGCTGTT	GGTTGTAGTT	42100
AACCCCTGAAT	TCCTTCCCAT	GTGCCCTTTG	CAAAGCCATC	AAGGCAGAGA	
GACTTGCCCTA	GCAAGTAGGA	TATTACAGTC	TTCTGTAATA	TAATCACATC	42200
CATGAAATCC	TCTATATATC	CCATCACCTT	TACCATATTC	TGTGGGTTAG	
AAACAAGTAG	CAGGTCCCTGC	CCACACTCGA	GAAGACCAGA	TGACACAAAG	42300
ATGTGATTCA	AAGTGGGGAT	CATCGGGGCC	ATCTTAGGTT	TGTCTGCAGT	
GATCACTGTG	CCATCTCTCT	CTCTCTCTTT	TTTTTTTTTT	TTTTTCCGAG	42400
ACGAAGTCGT	CACTCTGTCA	CCCAGGCTGG	AGTGCAGTGG	CATGATCTCA	
GCTTACCACA	ATCTCTGCC	CCCAGGTTCA	AATGATTCTT	CTGCCTCAGC	42500
CTCCTGAGTA	GCTGGGATTA	CAGGTGCCCC	CCACCACACC	CAGCTAATTT	
TTGTATTTTT	AGTAGAGACA	GAGTTTCACC	ATGTTGGCCA	GGCTGGTCTT	42600
GAACCTCTCA	CCTCAAGTGA	TCCACCCACT	TCGGCCTCCC	AAAGTGCTGG	
GATTACAGGC	ATGAGCCACC	ATGCCAGCC	CCATCTCTCT	TTAAAAACA	42700
AAACAACAAA	CAAAAACAT	AAAAAGAAGC	AGAGAACACA	TACACATCTG	
CATCTTCCCT	TGTTTACTTA	ACAATAGATC	TTGGAAGTCA	CTTCTCAGTA	42800
GAGGCTAGGT	TGGGCAGAGC	ATTGGATTCT	AGGCCAGTGA	GTTTGGACTT	
GACCATGGAG	ACACTAGGAA	GCCCATGAAG	GACAGAGAGA	GATGCCTCGA	42900
CCCTGCCAGT	CCTTTAGAAA	GATCACCAG	TGCTTTTTGT	ATACCAAACC	
CTATTTGAAA	TACTTACGTA	TATTAACCCA	TTTCCTTATC	ACCACAACCC	43000
TGCGGGAAGG	GAGATAGGCA	CTTTTATTAT	CTTCATTTTG	CAGATGAGGA	
CATTGAGGTC	CAGAGAGGTT	ATGTCACTTA	CTTAAGGTCA	CACAGCCAGG	43100

FIGURE 1E

AAGTGGTAGT	AGGGACTCTT	ACCCCTTGTTTT	TACAGATGAG	ATTGAATTAT	
CTCACGAAAA	CTCAGAAAAG	TTAAACAAC	TGCCTAAGTA	ACATACAGCT	43200
AATTAGTCGA	GGAGCCTGAC	GCATGTTGCT	CTAGCCTGGT	CACAGTTACA	
GAGGTGGCAA	GCAATGGCCT	GAACAGGACG	AACAACCAAA	TACCCAGGCT	43300
GGTGGCTCTT	AAACATGGTG	GGGTCAGCTA	ACGACAGCAA	CCAGGGTGGG	
CACGTGGTGCC	CCTCGCCCCC	GGCTGGTGCC	CTAACATCTC	CCTTTTCTCT	43400
ACCAGTTCAG	AATCTATAAC	GTGACCTACC	TAGAACCCTC	CCTCCGCATC	

A

[exon 7: 43406..

GCAGCCAGCA	CCCTGAAGTC	TGGGATTTCC	TACAGGGCAC	GGGTGAGGGC	43500
CTGGGCTCAG	TGCTATAACA	CCACCTGGAG	TGAGTGGAGC	CCCAGCACCA	
AGTGGCACAA	CTGTGAGTAT	CAAGAGGCCT	AAGCAATGGT	AATCTCCACT	43600

..43562]

CTCCATTCTT	CCCCGTGTCG	CAGACACTTC	CCCTGGCTGA	GTCTCTGGGC	
TTTTATATCA	TAGGATGCC	CTAATGGCAA	TCCTGCCATT	AGATACACCT	43700
GCTGTGGTGT	ATCTGCCAGG	TAGGCAGGCT	AGGCTGCAGT	AACACACAAG	

C

CCCACAATTT	CCATGGCTTA	ACACTATAGG	AATATATTTT	TTGCTCATGT	43800
AACAAGCTAA	CGTGAATGTT	GCTGGTTTGT	AGGTGGTTTC	CCTCCCTGTA	
GAAATCTGGG	GAGTGAGGTT	CTTTCCATCT	TGTGGTGCCA	TCATTCTCCA	43900
GGACAAAGAT	TCTTACCTAC	TTTTGTGTCC	TGGTTTCCTT	TGGCAGCCTG	
GTGAAGCCTA	TGGACCTCAT	TTCAGAAAT	TTTTAAATAC	ATAAAATCCC	44000
AGCCTGGGCA	ATATAGTGAA	ACCCCATCT	GTACAAAAT	TAGCCAGGCA	
TGGTGGCATG	CACCTGTAGT	CCCAGGTACT	GGGAAGGCTG	AGGTGGGAGG	44100
ATCACTTGAG	CCCAGGAGTT	TGAGGCTGCA	GTGAGCCGTG	ATCGTACCAC	
TTTACTCCCA	CCTGGGTGAC	AGAGCAAGAG	CCCATCTCTA	AAAATAAATA	44200
AATACAATGA	AATAAAATAA	AATAAATAGA	ACTACAGAGG	AAACTAATTG	
TATTGAAATG	CAGTTATAAA	ACATTTAAAC	ACATTTTTAA	TCTAGAGATA	44300
TATGTGCTTC	TTTATTAAGA	TCTATAAATA	ATAAGTTCTA	GGGGTAGCTC	
GCATAAATAC	TGTAATTTCA	AAGTAGATAA	GCATAAATAA	TACTTTATGA	44400
TACTGAAAT	GTGATGTGAT	ATGAGAATAG	CTGTGAGTTT	TGTTTTGCTG	
GGGACAGGAT	CACGTATGCT	GTCATTACTG	GGGTCTCTTC	CCTCCATTCT	44500
TTTTTTAAAA	TTGTATTTTA	TTTTATTTTT	AAAATTTTTAA	AATAAATAGA	
GACAGGGTAT	CACATATGTT	CCCAGGCTGC	TTTTGACCTC	CTGGGCTCCA	44600
GTGATCTTCC	CATCTTGGCT	TCCCAAAGTG	CTGGGATTAC	AAGTGGGAGC	
CAGTGTTCCT	GGCCCCCTCC	TCCATTCTTA	ATGGAAGGAG	ATGCTAGGTG	44700
TGAGAGGTTA	GGGAAAGTAA	AGATGTAATT	TCTTTCCCAT	CCAAGTTCTC	
AGACCCCTGA	ATTCTACCTG	CAGCCATGTT	GGTCCATCAA	CCCAAGTGTA	44800
AGAATCCCTG	CTCTAGGGCC	CCACCATTGT	CTGTATCCAG	CCAGCAGAA	
AGGCGTGATT	ATGGAGATCA	CATCTGCTTC	TTGAAAGCAG	ACAGCCCGGA	44900
AGTGGGCCGC	ATCACTTCCCT	CTCAAATTCT	ATTGGTGAAA	ATGGTCACAT	
GACTACACAT	AGCCACAAAG	GAGGCTGGGA	ACTTTCTCAC	TTGGAACCTA	45000
CATCCCAGAA	ACAACCTCTT	TCAGTGAGGT	ATCCCACAGG	TCTTTCCGAG	
TAGAAATATT	GATTATCTCA	CATAAAATGA	AGTCTTACAA	ATGGACCTAC	45100
TGGGTTTTGT	ACAGCAGCCA	AGTGATATCT	CTTCCCTTCT	GCTGTCTTCC	
CTTCTGCCGT	CCTTCACATG	GTGGCATTGT	ATCCTTAGAC	TTGCCACCCA	45200
TGCCCCCAGG	TTGGCCGTTG	CACACTGTCT	TACATAAAGC	AGGAAGGAAA	
GGAAAGGCTG	CTACGAGAGA	GTGTACCTTG	TGCATCTCTT	TTTTAATCAG	45300
GAAGCAAACA	TCTTTCTAGA	AGCTTCCCTA	GCAAAATTC	CCTTACATCT	
CATTGGCCAA	GACTGTTTACA	TGTTACATGG	TTACTGTTAT	TACTTGCTCA	45400
TTGCAAGGAA	GACTGGGAAC	TCAAATGCCT	GGAAAAAGGA	ACAGGATAAT	
CGTGATTGGC	TCAAGCCTTA	GGGTGGGCAT	GGCTCCCTGA	CAAGGGAGAG	45500
AGGAAAAAGC	TGTTGAGTGA	AGAAGACTGC	TTCAGTTTCC	CCATCTGTAT	
AATGGGAGGA	GTAAGGGCTG	TCGTGAAAC	TCAATGAAAG	AAGATTCTTC	45600
AACGTGGTAG	GTGCAGTGGC	AGCTGGCAGT	ACCCTGACCC	TGCCACCGCA	
CAGCCCTCTC	AGCATTGCTC	ATCCTGCACT	GTGGATATCA	GTTGAGCCAC	45700
GTGTCTCCTG	CCCTGGGCTG	TGAGCTCCAT	AGGCAGGGTC	TCCATGGCTG	

FIGURE 1F



TATCTCCAGA	ACCCAGCACA	GAACCAGGTG	CTTGGGAAAG	TTTTGAAATTG	45800
ATTCTCATCT	GCCATTGGCA	TGGGGAAGGG	AACTAGCTTG	TATGAAACAG	
ATAACAATGT	ATGGGACCCT	CATTCAATTAT	TTCAGCAAAT	ATTTGCTGAG	45900
TTCCTCCTAC	ATGGCTAGCC	CTGTGCTAGA	CACTGGGGAA	TCGGCGATGA	
ACAAAGCAGA	TAGAAATCCC	CACTCTTGTG	GAGCTGACAT	TCTGGAGGGA	46000
GAGACAAAA	GCAAACATAT	AAAGAAAGAA	AGAAATCACA	TGGATCTGGA	
TGACAGTGAG	TGCTGGGAAG	AAAATAAAAAG	CAGAGGAAGG	GGATGGAGCG	46100
ATGGGCAGGG	GGCAACGGTA	GGGAGGGTGT	CGGGGAAAAC	TTTTTGGAGA	
ATGTGACGAT	GAAAGTGAAC	AAGGAGAAGT	CAACCGTGT	GAGATGATGG	46200
CAGCTAATGA	TGTGGACAGG	CCACTCTGTT	CTGAGTGCAT	TATCTATTGA	
TTCATCATGT	CATCCTCGCA	ACAGCCCTGC	ACGATCAATT	CTGTCATTAA	46300
CCCCATAGTA	CAGATGAGGA	TGCGGAGGCA	CAGAGAAGAT	AAGGGACTTG	
TCCTGTGTCA	CACAGCAAGG	AGCCATCCGG	CTCCTAAGTT	GGTGCATTTG	46400
ACTTCTGTGC	TTCCGGAAAG	AAAGAGCAGC	AAGTTTAAGA	TCTGGAGGTG	
GCATGAGCT	TTGGAGGAGC	AGGGGGCAAT	GAGGTGGCCG	GTGTGACGAG	46500
GACTCAATGT	GCAAGAGGGA	GAGTGGTGGG	GAGATGAGGT	GGAGGGGTGG	
TCGGCGGTCA	GATCGTGGAG	GGTCTCGGAC	GAGGGTCCTG	ACCCTGGGTC	46600
TCCAGTCTTG	GGAAGTGGAG	CCCAGGCTGT	ACCATGGCTG	ACCTCAGCTC	
ATGGCTTCCC	CTCCCACTTC	CAGCCTACAG	GGAGCCCTTC	GAGCAGCACC	46700
[exon 8: 46674..					
TCCTGCTGGG	CGTCAGCGTT	TCCTGCATTG	TCATCCTGGC	CGTCTGCCTG	
TTGTGCTATG	TCAGCATCAC	CAAGTGAGTC	CTGGGCCCGAG	TGCTGCCGAG	46800
..46773]					
CAGTCCCTCT	GGAGTGCAGG	GTGGCAGGGA	CTTGCCCCCTC	TAGTCTGCCC	
CTTTGCAGTC	CTCTCAGTCA	ATAATACGTA	TTTACTGAGC	AGCTACTACA	46900
CACCTTGAGA	GTAGAGCTGA	GAACATATCG	ACAAGGACCC	CACTTTTTTC	
TTTTTTTTCTT	TTTTTTTTTTT	TTTTGAGACG	GAGTCTCACT	CTGTCACCCA	47000
GGCTGGAGTA	TAGTGGCACA	ATCTTGCCCTA	ACAGTAACCT	CCGCCTCCCG	
GGTTCAAGCA	ATTCTTCTGC	CTCAGCCTCC	AGAGTAGCTG	GGATTACAGG	47100
CGCATGCCAC	TATGCCCGGC	TAATTTTTTTG	TATTTTTGGT	AGAGATGGGG	
TTTCACCATG	TTGGTCAGGC	TGGTCTCGAA	CTCCTGACCT	CATGATCTGC	47200
CTGCCTCAGC	TGCCCAAAGT	GCTGGGATTA	CAGGTGTGAG	CCACTGCACC	
CAACCAGGAC	TCCACATTTT	TAAAACCGGC	ATCCTACTGG	GGAGACTGAA	47300
AATACATATC	AATCACAAC	AGGTGGTTTTT	CCATAGTGAC	CCACTCTCTG	
AATGCACTAG	ACCAGGGTGG	AGGCCAGAGA	TCTTCTGGGG	TGCTTTTTTGC	47400
AAGGGGGACC	AGGATAAGGC	TCTCCAAGGA	GGGAAAATTT	GAGGGGGGCC	
CTGACTGGGG	AGAATGAGCT	GGCCAGGGAT	AAGCAAGATG	GAGTCATCCC	47500
ACATCCCCTT	ACAACACTGG	GTGCCTGGGC	AACTGGGGGC	ATTTGGGGGC	
ATGTGGTAGG	AGCCAGAGGA	ATTTGCGACG	ATTGCCCTGA	TGGAGTCAGG	47600
AGACCTGGGT	TTGAATCCTG	GCCTTGGAGC	TTGGTAGCTG	GCGGCCGACA	
AGTTGCTGAA	ACCCCTGAGC	CTGGGGTTCC	TGCTTTGCAG	AGTGACAGTG	47700
ATGGTGAGAA	CATATTTTCT	CAGCCAGAAG	AGGCCAAATC	ACAGTAAAGG	
CTGAGGGAGG	AGATGAGTGG	CGAGTGGCTG	GGAGGTGGTG	GAAGGAGCCT	47800
CGTTTCCAGA	GAGCTCTTGC	CAGCCCTTGG	AATCATGGTG	TCTCAGAGCC	
TCAGTCCTCC	CATCTCTGAA	ATGGGACTAG	CAAGCTCAAC	CTCACTAAGT	47900
CAGGATTAGA	GGTGGCTAAG	GATTATTAAC	ATGATTGATG	AAAGTGCCCA	
CTCTTGGCCC	AGCACACACT	AGGTAGGCAG	GGAATGCAAA	TTCCCTCCA	48000
TATCTTGTCA	CTGATGCCTC	CGAGCAACCT	TGGACTGATC	GCCTTGCTCT	
GAGCCTCAGT	TTCCCCATCA	CCTGTACCTC	TTCCCACTCC	CCATCACTAT	48100
ATCCAGCAT	GCCAGCCTCT	TTGCTGTTCT	TTGTCTTTGG	TTTCTTGTTT	
TGTTCTGTTT	TTTAGACAGG	GTCTCACTCT	GTTAGCCAGG	CTGAAGTGCA	48200
GTGGCGCGGT	TACGGCTCAC	TGCAGCCTCC	AATTCCTGGG	CTAAAGAGAT	
CCTCCCATTT	CAACTTCCAG	AGCAGCTGGG	ACAACAGGCG	CTTGCCACCA	48300
CACCTGGCTA	ATTTTCTTAT	TTTAATTTAA	TTTTATTTTA	TTTTTTGGGA	
CAGAGTGGAG	TCTCAAAAAC	CAAGCTAGAG	TGCAGTGGTG	CGATCTCGAC	48400
TCACTGCAAT	CTCTGCCTCC	CGGGTTCAAG	CGATTCTCCT	GCCTTAGCCT	
CCCGACTAGC	TGGGATTACA	GGCGTGTGCC	ACGACACCCA	GCTAATTTTT	48500

FIGURE 1G

GTATTTTTAG	TAGAGATGGG	GTTTCACCAT	GTTGGCCAGG	ATGGTCTTGA	
ACTCCTGACC	TCAAGTGATC	CACCCACCTC	GTTCTCCCAA	GGTGCTGGGT	48600
ACAGGCATGA	GCCACTGTGC	CTGGCCAATT	TTCTTACATT	TTGTAGAGAC	
TGGCTGTCAC	TTATGTAGCC	CAGGCTGATC	TTGAACTTCT	ACCCCTTTAT	48700
CTTTATTCAT	GGCACTTATT	ACCATGAATG	AATGACCTCA	TATAAGCATT	
CTTTTCGTTT	TTTTTTTTTT	TTCTTTGAGA	TGGAGTCTCA	TGTTGTCCCC	48800
CAGGCTGGAG	TGCAGTGGCG	CGATCTCAGC	TCACTGCAAC	CTCCGCCTTC	
CGGGTTCAAG	CGATTCTCCT	GCCTCAGCCT	CCTGAGTAGC	TGGGATTGCA	48900
GGCGCCTGCC	ACCATGCCTG	GCTAAGTTTT	GCATTTTTAG	TAGAGACGGT	
GTTTCACCAT	ATTGGCCAGG	CTGGTCTCGA	ACTTCTGACC	TCAGGTGATA	49000
CACCTGCCTT	GGCCTCCCAA	AGTGCTGGGA	TTACAGGCGT	GAGCCGCCAT	
GCCTGGCCTC	ATATAAGCAT	TTCTGTCTCC	ATTTATCATC	CATCTTTCCC	49100
TCTTGAAGGT	CAGTTTCACC	AAGGCAGGCA	TCTTTGTCTC	GTTCACGTGT	
GTAGCCTCAG	GGCCAGGCAC	AGTGAGTCAA	ACATAGAAGG	TGCTCAATAA	49200
ATATGTGTTT	ATTTATTGAA	ACCATGGGCA	GAGGCTAATT	CAGAAGCGGT	
CTGAGGACCT	TACCTCCCAG	TGATGATGCA	CCATGGCCCC	AGGCAGGCCA	49300
GGAAGAGAGA	AGGGTTGTGT	TTCTCCGTAG	GTCCCCCAGC	TTCCCAGGCC	
ATCCCAGGCC	ATTCCTTGGT	CATTTGCCCT	CAGCTGCTCT	GAAAAAGGGA	49400
TTGTTGAGGG	GAACCTAGAA	TCCTCTCTCT	GCAGTTTGAG	TCTTTCCTAA	
TCCCCTGGGG	TCTCATTCCT	ACTGAGGACA	TAGGTGGCCT	CCTCAGGAAC	49500
TCTGTGCTGG	GTAACAGAAT	GCGGGAGTGT	GAACCTGGCT	CTGCCACCTA	
CCAGCTGTCA	CTCCACCTCC	TTGGGCCTCA	CTCTCCTCAT	CTGTAGAATA	49600
GGGTTAGCAA	TAGAATCCAT	GTCACCAGGT	TAGAATGATG	AGTCAGTGGT	
TTGACCTCCA	GAAACTAATC	AGCCTGATCT	CTGATGCCAA	ATAAGTATTG	49700
GTGATAACGA	CCACTTTTAT	GGGAGGAGCG	TTCACTTGTC	AATAATTCAG	
AGATCAACAC	CTTTTCCTTT	TGTTTTTCAG	GATTAAGAAA	GAATGGTGGG	49800
[exon 9: 49781..					
ATCAGATTCC	CAACCCAGCC	CGCAGCCGCC	TCGTGGCTAT	AATAATCCAG	
GATGCTCAGG	TAGGAGTAGG	CGTGGATGAG	GACATGTGGG	ACTGTGTACA	49900
..49859]					
TGAAGAAGTG	TGGTTCAGAA	CACCTGGGCT	GTTAAGGACC	TTCACCTGGCT	
TCTGGAATGG	CAAATAGACA	GTCAGGAGGG	TTGCAGGGGA	GACAGAGGCA	50000
GAAGCCGAAT	GAGGTCATTA	GCAGACCAGA	GGCTTTCCCG	CCCTTCCCTT	
TGGCAATCCC	AGCCTGGGGT	GGGCTTCTCT	GGGGTTGGTT	TCCTGTTTTT	50100
TTCCCTCCCC	TTGGGAGAAT	GACCCTTGGG	TCATCATCAC	TGTGTCAATC	
CCTGGGGAGG	TGCCAGTACC	AGGGCTAGAG	GCCAGAAGGA	GTGGAGGAAG	50200
GAGAGGGTGA	CAGGCTTTCT	GTGTCTTCTT	CTTAAGCATA	GGAAACTGCC	
CCCGAAGCAC	TAGCAAATCC	CTTCCGGGTT	CTCATTTGGC	TGAAATGTAT	50300
CCCACCCCTA	AGCCAGGGGT	GGAGTCAGCT	TCCCCAAGGC	GATGGTCCTG	
TGGGTGAGTG	GGTGGGGTTT	GCCTGAGCAA	GATGAGAGTT	CTCTAGGTAG	50400
GAGAAAGGGG	GATTATAGGT	CCTGTCTAGA	AGAGAAGGTC	TGAGGGTCCT	
TGCTTTTCCA	GGGACTCTGG	AATCTAGTGT	TGGCTTTGAA	TCCTGACTCT	50500
GCCACTCACT	GGCAGTGTGG	ACTTGAGCAA	GTTGCTTAAT	TCTCTGAGCC	
TCAGTTTCCT	CTTGTGGGTT	ATAACAGTGT	TTACCTGGTA	GGACAGATAT	50600
TGGAATTTAT	TGAGACAATA	CATATAAAGT	GCATATTCCA	GCCTCTTGCA	
AATACCAAGT	GCCATTTATG	TATCAGTTAG	TGTTTGCTGT	GTAACAAATG	50700
ACCCCGAAAT	GTAGAGGGTT	ACAACAACCT	TATTTAGCTT	ATGCTTCTGC	
AGGCTGGCAT	TTGGGGCTGG	GCTCAGCAGT	GAGGGTGGCG	GGGGAGGCTG	50800
GGCTGGGCTG	GGCTGGGCAG	ATCTGAATTG	AGCTGACCCG	TCCCCGTAGC	
CTCCCTCCGT	GTCTGACAGT	TGGCTTTTTT	TTTTTTTCTC	TTTTTCTGAG	50900
ACGGAGTTTT	GCTCTTATTG	CCCAGGAGTG	CAATGGCATG	ATCTTGCGTC	
ACTGCAACCT	CTGCCCTCCT	GGTTCAAGCA	ATTTTCTTGC	CTCAGCCTCC	51000
CAAGTAGCTG	GGATTACAGG	CATGTGCCAC	CACGCCAGGC	TAATTTTGTA	
TTTTTAATAG	AGATGGGGTT	TCTTCATGTT	GGTCAGGCTG	GTCTGGAAC	51100
CCTAATATCA	GATGATCCAC	CCACCTCAGC	CTCCCAAAGT	GCTGGGATTA	
CAGGCGTGAG	CCACTGCACC	CAGCCTAGTT	GGCTGACTTT	TACCTGGGAC	51200
AGTGCAAGTG	CCTGAGCCAT	GTGCCCTCTCA	CTCTCCAGCA	GGCCGGCCCA	

FIGURE 1H

GGCTTGTTTA	CAGAGTGGCT	CAGTTTTCAA	GGGTGGGAAG	TCCCAAGGCT	51300
TCTTGAGGCC	TAGGCGCAGC	ACTGGCATGA	TATCACTTCC	ATCACATTCT	
ATGGGCCCCA	GCAAGTCCCA	GGGCCAGTGT	AGATTCAAGG	GATGGGAGGA	51400
GATTCAGAGC	ACTCCTCTGT	GGCCACTTTT	GCCATCGACC	ACAGTCCCTG	
TAAATATTAG	GACAATGTAA	TTAATTCCCA	GGAATCTGAG	GCTCAGAAAG	51500
CGTAAGTGAC	CTGTTGGACT	TCTGATCTGT	GTGATGTCGA	GGCTTGTACC	
CCTTCCTGAG	CATTGCCGTA	CTCCAGGCCG	GGCTGCAAGG	CCACTCTGCT	51600
CTTTCATTGG	CTGTCTCTGT	ATTTTAGGGG	TCACAGTGGG	AGAAGCGGTC	
[exon 10: 51628..					
CCGAGGCCAG	GAACCAGCCA	AGTGCCCGTA	TGTATCTGAA	CTTAGGTCAC	51700
..51677]					
AGCCTGCATG	CATTGGGAAG	GTGATAGAAT	TGGAGAGGCA	AGCCCCTAGC	
TCCATGTCTG	CCTTCTCTTC	CCTGCATTCT	GTAATTGCCC	TGTGACATTA	51800
GCCTTCAAGG	GACGGCAGGA	GGAGGGGTGT	TCTGGAACG	TGGACTGCTG	
GCCAAGCCCC	CTGAGTTTCA	CTGGTGTGTC	AGGTACATGG	TGATACCCCT	51900
TGGGAGTGGCT	GTTATAGTTA	ACAACCAGAG	CAGCCGTGCC	TGTTGTTAAA	
ATCTTGACCT	AATTGTATAC	TTGTGCGCAA	ATAGCCACTA	TCCTGAACAC	52000
TCCCCTCCTT	TTTTTTAATA	TACAGGATCT	CACTCTGTGG	CCCAGGCTGG	
TGTGCAGTGG	TGCGATCATA	GCTCACTGCA	CCTTCAAAC	CCTGAGCTCA	52100
AGTGATCCTC	CCATCTTAGC	CTCCCGAGTA	GCTGATACTA	CAGATGTGCA	
TTACCACGCC	TGGCTATTTT	AAAAGGTTTT	TGCCGTGTAAT	TCCAGCTACT	52200
CAGGAGGCTG	AGGCATGAGA	ATCACTTGAA	CCCGGGAGGC	AGAGGTTGCA	
GTGAGCGCAG	ATTGTGCCAC	TGCACTCCAG	CCTGGGCGAC	AGAGTGAGAC	52300
TCTTGTCTCA	AAAAAAATAA	TACCAAAAAA	AGTTTTTGTA	AAGACAAGCT	
CTCGCTGTGT	TGCCCCGCCA	CTGTGGCCTC	CTTAGCTTCT	TCCCTGGGGC	52400
CTGCTGGACC	TTTCCATACT	CCAGAAACTA	AAGGGGGTCC	AGGACCCTGC	
TTCAACCCTA	GGATCCCGCA	TCTTTTTTTT	TTTTTTTTTT	TTTTGGACGC	52500
AGGGTCTTGC	TGTGTCCCTC	AGGCTGGAGT	GCAGTGATTC	ACTGCAGCCT	
CAAACCTCGT	GGCTCAAGTG	ATTCTCTAGC	CTCAGCCTTC	TAAGTAGCTG	52600
GGACTACAGT	CATACACCAA	CATGCCCAGC	TAATTTTCCT	TTTTTTTAAT	
TCTTGTAGAG	ATGTTTGAGA	CGGCTTGGGC	TCTGTTGCCC	AGGCTGTTCT	52700
CAAACCTCCT	AGCTCAAGCG	ATCCTCCCTC	CTCAGCCTCC	TAAAGTGCTG	
GGATTACAGG	CGTGAGCCAC	CGCACCCGGC	TTCCATATCC	TTTCTAATTG	52800
GTCAATGGCTT	GGGATAATGG	TGTTGCTTTT	AATTATCATC	ATCCATAAAG	
ACTTTTTCTT	ACTCAACAGA	TCTGAGCTTG	TATTTGGTGC	CCAGGACATG	52900
TGCTGGGTTT	CCGAAATCCC	AAAGACACAG	ACCCTACCCT	CAGGGATTTC	
TCATTCTAGC	AACATAGACT	GATCAATTAC	TGATTATAAC	GTTAGAAGGC	53000
ATGTCTGAAG	TAGACAGCCA	TCAGGACATG	GTGATTTCAG	GCTGGGCTTT	
C					
GAAGAATGAA	TAGGAGTTTT	TCAAGTGTCT	AAACTGAACC	CTGACCAACC	53100
T					
TTTGCTTTTG	CAGACACTGG	AAGAATTGTC	TTACCAAGCT	CTTGCCCTGT	
[exon 11: 53114..					
TTTCTGGAGC	ACAACATGAA	AAGGGATGAA	GATCCTCACA	AGGCTGCCAA	53200
C					
AGAGATGCCT	TTCCAGGGCT	CTGGAAAATC	AGCATGGTGC	CCAGTGGAGA	
TCAGCAAGAC	AGTCCCTCTG	CCAGAGAGCA	TCAGCGTGGT	GCGATGTGTG	53300
GAGTTGTTTG	AGGCCCCGGT	GGAGTGTGAG	GAGGAGGAGG	AGGTAGAGGA	
AGAAAAAGGG	AGCTTCTGTG	CATCGCCTGA	GAGCAGCAGG	GATGACTTCC	53400
AGGAGGGAAG	GGAGGGCATT	GTGGCCCGGC	TAACAGAGAG	CCTGTTCTCT	
C					
GACCTGCTCG	GAGAGGAGAA	TGGGGGCTTT	TGCCAGCAGG	ACATGGGGGA	53500
T					
GTCATGCCTT	CTTCCACCTT	CGGGAAGTAC	GAGTGCTCAC	ATGCCCTGGG	
C T C					
ATGAGTTCCC	AAGTGCAGGG	CCCAAGGAGG	CACCTCCCTG	GGGCAAGGAG	53600

FIGURE 11

CAGCCTCTCC	ACCTGGAGCC	AAGTCCTCCT	GCCAGCCCGA	CCCAGAGTCC	
AGACAACTG	ACTTGACAG	AGACGCCCT	CGTCATCGCA	GGCAACCCTG	53700
CTTACCGCAG	CTTCAGCAAC	TCCCTGAGCC	AGTCACCGTG	TCCCAGAGAG	
		C			
CTGGGTCCAG	ACCCACTGCT	GGCCAGACAC	CTGGAGGAAG	TAGAACCCGA	53800
GATGCCCTGT	GTCCCCCAGC	TCTCTGAGCC	AACCAC'TGTG	CCCCAACCTG	
AGCCAGAAAC	CTGGGAGCAG	ATCCTCCGCC	GAAATGTCTT	CCAGCATGGG	53900
GCAGCTGCAG	CCCCCGTCTC	GGCCCCCACC	AGTGGCTATC	AGGAGTTTGT	
	T		G	A	
ACATGCGGTG	GAGCAGGGTG	GCACCCAGGC	CAGTGCGGTG	GTGGGCTTGG	54000
GTCCCCCAGG	AGAGGCTGGT	TACAAGGCCT	TCTCAAGCCT	GCTTGCCAGC	
AGTGCTGTGT	CCCCAGAGAA	ATGTGGGTTT	GGGGCTAGCA	GTGGGGAAGA	54100
GGGTATAAG	CCTTTCCAAG	ACCTCATTCC	TGGCTGCCCT	GGGGACCCTG	
CCCCAGTCCC	TGTCCCTTGG	TTCACCTTTG	GACTGGACAG	GGAGCCACCT	54200
CGCAGTCCGC	AGAGCTCACA	TCTCCAAGC	AGCTCCCCAG	AGCACCTGGG	
		T			
TCTGGAGCCG	GGGGAAAAGG	TAGAGGACAT	GCCAAAGCCC	CCACTTCCCC	54300
AGGAGCAGGC	CACAGACCCC	CTTGTGGACA	GCCTGGGCAG	TGGCATTTGTC	
TACTCAGCCC	TTACCTGCCA	CCTGTGCGGC	CACCTGAAAC	AGTGTCTATG	54400
CCAGGAGGAT	GGTGGCCAGA	CCCCTGTCT	GGCCAGTCCT	TGCTGTGGCT	
GCTGCTGTGG	AGACAGGTCC	TCGCCCCCTA	CAACCCCTT	GAGGGCCCCA	54500
	G				
GACCCCTCTC	CAGGTGGGGT	TCCACTGGAG	GCCAGTCTGT	GTCCGGCCTC	
CCTGGCACCC	TCGGGCATCT	CAGAGAAGAG	TAAATCCTCA	TCATCCTTCC	54600
ATCTGCCCC	TGGCAATGCT	CAGAGCTCAA	GCCAGACCCC	CAAATCGTG	
	C				
AACTTTGTCT	CCGTGGGACC	CACATACATG	AGGGTCTCTT	AGGTGCATGT	54700
				C C	
	..54692				
CCTCTTGTG	CTGAGTCTGC	AGATGAGGAC	TAGGGCTTAT	CCATGCCTGG	
				T	
GAAATGCCAC	CTCCTGGAAG	GCAGCCAGGC	TGGCAGATTT	CCAAAAGACT	54800
		G			
TGAAGAACCA	TGGTATGAAG	GTGATTGGCC	CCACTGACGT	TGGCCTAACA	
CTGGGCTGCA	GAGACTGGAC	CCCGCCAGC	ATTGGGCTGG	GCTCGCCACA	54900
TCCCATGAGA	GTAGAGGGCA	CTGGGTGCCC	GTGCCCCACG	GCAGGCCCTT	
GCAGGAAAAC	TGAGGCCCTT	GGGCACCTCG	ACTTGTGAAC	GAGTTGTTGG	55000
CTGCTCCCTC	CACAGCTTCT	GCAGCAGACT	GTCCCTGTTG	TAACTGCCCA	
AGGCATGTTT	TGCCCACCAG	ATCATGGCCC	ACATGGAGGC	CCACCTGCCT	55100
		G			
CTGTCTCACT	GAAGTAGAAG	CCGAGCCTAG	AAACTAACAC	AGCCATCAAG	
				A	
GGAATGACTT	GGGCGGCCTT	GGGAAATCGA	TGAGAAATTG	AACTTCAGGG	55200
AGGGTGGTCA	TTGCCTAGAG	GTGCTCATTC	ATTTAACAGA	GCTTCCTTAG	
GTTGATGCTG	GAGGCAGAAT	CCCGGCTGTC	AAGGGGTGTT	CAGTTAAGGG	55300
GAGCAACAGA	GGACATGAAA	AATTGCTGTG	ACTAAAGCAG	GGACAATTG	
		A			
CTGCCAAACA	CCCATGCCCA	GCTGTATGGC	TGGGGGCTCC	TCGTATGCAT	55400
GGAACCCCA	GAATAAATAT	GCTCAGCCAC	CCTGTGGGCC	GGGCAATCCA	
		T			
GACAGCAGGC	ATAAGGCACC	AGTTACCCTG	CATGTTGGCC	CAGACCTCAG	55500
GTGCTAGGGA	AGGCGGGAAC	CTTGGGTTGA	GTAATGCTCG	TCTGTGTGTT	
			T		
TTAGTTTCAT	CACCTGTTAT	CTGTGTTTGC	TGAGGAGAGT	GGAACAGAAG	55600
GGGTGGAGTT	TTGTATAAAT	AAAGTTTCTT	TGTCTCTTTA	TTTTTTATGT	
ATTAACCAAA	CATACCTCCA	GACACTGCTG	TGAGTGCTGT	GTCTCTGTTA	55700
ACTCCTGGAA	TTCACCCATC	CAGAGGAACC	AGGATGCAAG	AGGTTAAGAA	

FIGURE 1J

ACTTGCCGTC	TGGGTTTGGG	TTCCCCATAC	AAGGATTCAA	ATAGTTGATT	55800
A					
TAGGAAGTAA	TCCCAGGAAA	CCCTGCTAAG	GTAAGTGGGA	ACTGAGGCAG	55900
GGAAGGACAC	AAACCAAGAA	AGTGTTACCT	GAAAGGGGTC	CAGATGCAGA	
CCCCAAAAGA	GGGTTCTTGA	ATCTCATGCA	AGAAAGAATT	CAGAGCGAGT	
CCATAGAGTC	AGTGAAAGCA	AGTTAATGAG	GAAAGTAAAG	GAATAAAAAGA	56000
ATGGCTACTC	CGTAGACAGA	GCAGCCCTGA	GGGTTGCTGG	CTGCCTATTT	56100
TTATGGTTAT	TGATTAATTA	TATTCCAAAC	AAGGGGTGGA	TTATTATGCC	
TCCCTTTTAG	ACCATATAGG	GTAACCTCCT	GATGTTGCCA	TGGCATTGTG	
AAACTGTCAT	GGCGCTGTTG	GGAGTGTAGC	AGTGAGGACA	ACCAGAGGTC	56200
ACTCTTGTTG	CCATCTTGGT	TTTGGTGGGT	TAGAGCCATC	TTCTTTACTG	56300
CAACCTGTTT	TATCAGCAAG	GTCTTTATGA	CTTGATATCGG	TGACGACCTC	
CTGTCTCATT	CTATGACTAA	GAATGCCCTA	ACCTCCAGG	AATGCAGCCC	
AGTAAGTCTC	AGCCTCATT	TACCCAGCCC	CTCTCAAAG	CTCCAGTTTA	56400
AATAAACCTC	TGACAAAAGG	GTGAGTTATT	CAACAGATTA	CCAGCATGAG	56500
TAACCTGATG	TTACCTGCCG	GGGATCTCTG	GAAGACCATG	CATGGCACAT	
GCCCAGTTAT	GCCTGCAAAG	GAGAGGGAGC	TGGGGTATTT	GTCCACCAGC	
TCCCATCTGT	CATTGGCTGA	GAGCTGCTTC	CAGGAGCATT	AATTCTCCAG	56600
CACTTCCAGC	TACTCCAGGA	AAAAAAAAAT	TCTTCAACTG	AGAGTTGGAG	56700
GTGTTGAGAG	ACTCTGGCAC	ACCAAGAAGA	CAGGAACAGG	ACACCAACAG	
TGGCTGATGA	TACACTGCCA	AGGTCACACA	GCTAGTTAGC	AACAGATCTA	
TAGTGGAATC	CAGACAGTGT	CTCCATCACC	CAGGCTCTCT	GTAAGTATCT	56800
GCGCTTCACA	TCCGAGGCAG	GCAGAGGGAT	GGTGTGGGCC	TTAGATGGGA	56900
AGGCTGGGAA	CCTGAAGCTC	CTATGTCTGT	ATCACTTTTG	CTTCTCTGAG	
TAGCTGCCCT	GATTTCACAC	TTGAGGGGCT	TGGCCATTTT	AGATTCTCTC	
CTGCTCTAGG	AGCCTACATA	CTACACTGGA	AATGATGGGG	AGCTCTCTAC	57000
CTCACATGCA	GCCTGATGTT	TGTTAGAAAC	ACCTCCTTGC	GCCAGGCATG	57100
ATGGCTCATG	GCTGTAATCC	CAGCAATTTG	GGAGGCTGAG	GCGGGTGTAT	
CACTTGAGGT	GAGGAGTTCA	AGACCAGCCT	GGCCAATATG	GTGAAACCCCT	
ATCTCTACCA	AAAAATAAAA	AATTAGCCGG	GTGTGGTGGT	GGGTGCTCTG	57200
AATCTCAGCT	ACTTGGGAGG	CTGAGTTGGT	AGAATTGCTT	CAACCTGGGA	57300
CGCGGAGGTT	GCAGTGAGCT	GAGATTGTGC	CATTGCACTC	CAGCTTGGAT	
GACAGAGTGA	GACCTGTCT	CAGGAAAAAA	AAAAAAAAC	AAAAAAAACC	
TTGTTCTAAG	CCAAAATCAA	TCCCTTTAGC	TGCCCAAATC	ACACAGTTTA	57400
CAGATGGAGA	AACAGTTTTA	GAGAGGAAAA	GGGACTTGCC	CAAAGTCACC	57500
CAGAGAATGG	CAGAGCCTGA	ACTAGCCTTC	TGGACTTCTT	GCCTCCAAAA	
GCTCTTTATA	ATAAAATATA	ATTTTAAATA	AAAAATAGTTA	TCTGTTTAGG	
GCCAAGCAAT	ATGCTAAGTG	CCGTCCAGCC	ACTGTGTCTAT	TTACGTCTCC	57600
AAACAGCTCT	AGTTGGGAGG	CTCAATGATT	ATCCCAATTT	TACAGATAAG	57700
GAACAGGTC	CAGAGAGGTT	GAGGATTAGC	CTAGAACCAC	ACAGCTAGGA	
AATCCTGGAG	CCAGGATTTG	AACCCGGGTC	TGACCTAAGA	GCTCCCAGCC	
GCCGTGATAT	ATCAGCTTAT	GTCTATCCTGA	CACCTACGCA	GATGTCGGCT	57800
CGAATCCACT	TTGCCTGAGC	ATTGTCTCAG	AGAAATCTAA	TTTAAAAATT	57900
AGGCAGCAAA	TAGAAAATAT	ATTTGACTGC	TAGAGATGCA	ATGGGACTGG	
GAGCCCAACA	AAGGATCTTA	GGCAAAAGAA	ATCCAAGTTG	TTGGCCTCAG	
CAACTATTAC	TGAAGTGGCT	GGGCTTTGGG	AAGCTACAGA	GGGATGAGAA	58000
GACCTGGTGG	ATCAGGTGGG	CCCAACTCAG	GCTGGCCCCC	ACCCTGCAGG	58100
AAGTAGGAAA	AGTCCAGGGT	CATAGGCCCA	GTGAGATGCC	GGCTGCGGGA	
GTTTCAGCCT	CCGGGGCTGG	ACCAGAGGGC	AGGAGGGGAC	GCCCCTGGGT	
AGCAGCGCCA	GAGTGGGCTG	AGTGGCCTGG	GCCCCTGCGG	GGGAGCTTTC	58200
AGAGATGTTG	ATTTGGGGGT	ACTCCCTCAG	CCCTGCCTTT	ACACAGAATT	58300
TGTGGGGGAT	GAGGGGAGGG	GGAAAGGGGG	GAGGAAGGCA	GTGAGTGCAT	
CTGAATTTTT	TTTTTTTTTT	TACAAAAAGT	GGCTTATTGC	ATTTTCTCTGA	
TTACTCTATC	AGCACGTGCA	GACCTTTTCC	TATTCAGAGA	AAGCCTGAAG	58400
ATATAAAGAG	GAAAGTGAAG	AAAAACCACC	GGAAATCCCA	TCCCCGCCCC	58500
AGCATCTGGC	ACTGTGTGGG	CGATCACGAA	ATGAGCGCTT	GTTTTTGAAG	
GCGTAGTATC	TCCGTGAACA	TCCGGTTGAA	CAACCTTTCT	GACTTTATTT	

FIGURE 1K

WO 01/04270

PCT/US00/19094

TTCCCACGAA	AGTTATTAAT	TAAAAAACAA	AAAGCAAAAC	ACCGAAAAAA	58600
CAAAAAACCC	AGCAAGTGTT	TGAGCTCCCA	CCACGAGGGA	GGCCTGACGT	
CACTGGATCC	TCCCGGCAGC	CGATGAGGCT	GCATGGGACT		58690

FIGURE 1L

POLYMORPHISMS IN THE CODING SEQUENCE OF IL4R $\alpha$ 

ATGGGGTGGC	TTTGCTCTGG	GCTCCTGTTC	CCTGTGAGCT	GCCTGGTCCT	
GCTGCAGGTG	GCAAGCTCTG	GGAACATGAA	GGTCTTGCA	GAGCCCACCT	100
GCGTCTCCGA	CTACATGAGC	ATCTCTACTT	GCGAGTGGAA	GATGAATGGT	
CCCACCAATT	GCAGCACCGA	GCTCCGCCTG	TTGTACCAGC	TGGTTTTTCT	200
GCTCTCCGAA	GCCCACACGT	GTATCCCTGA	GAACAACGGA	GGCGCGGGGT	
		G	T	A	
GCGTGTGCCA	CCTGCTCATG	GATGACGTGG	TCAGTGCGGA	TAACATATA	300
				C	
CTGGACCTGT	GGGCTGGGCA	GCAGCTGCTG	TGGAAGGGCT	CCTTCAAGCC	
CAGCGAGCAT	GTGAAACCCA	GGGCCCCAGG	AAACCTGACA	GTTCAACACCA	400
ATGTCTCCGA	CACTCTGCTG	CTGACCTGGA	GCAACCCGTA	TCCCCCTGAC	
AATTACCTGT	ATAATCATCT	CACCTATGCA	GTCAACATTT	GGAGTGAAAA	500
CGACCCGGCA	GATTTTCAGAA	TCTATAACGT	GACCTACCTA	GAACCCTCCC	
T					
TCCGCATCGC	AGCCAGCACC	CTGAAGTCTG	GGATTTTCCTA	CAGGGCACGG	600
A					
GTGAGGGCCT	GGGCTCAGTG	CTATAACACC	ACCTGGAGTG	AGTGGAGCCC	
CAGCACCAAG	TGGCACAAC	CCTACAGGGA	GCCCTTCGAG	CAGCACCTCC	700
TGCTGGGCGT	CAGCGTTTCC	TGCATTGTCA	TCCTGGCCGT	CTGCCCTGTTG	
TGCTATGTCA	GCATCACCAA	GATTAAGAAA	GAATGGTGGG	ATCAGATTCC	800
CAACCCAGCC	CGCAGCCGCC	TGCTGGCTAT	AAATAATCCAG	GATGCTCAGG	
GGTCACAGTG	GGAGAAGCGG	TCCCGAGGCC	AGGAACCAGC	CAAGTGCCCA	900
CACCTGGAAGA	ATTGTCTTAC	CAAGCTCTTG	CCCTGTTTTTC	TGGAGCACAA	
			C		
CATGAAAAGG	GATGAAGATC	CTCACAAGGC	TGCCAAAGAG	ATGCCTTTCC	1000
AGGGCTCTGG	AAAATCAGCA	TGGTGCCCGAG	TGGAGATCAG	CAAGACAGTC	
CTCTGGCCAG	AGAGCATCAG	CGTGGTGCCA	TGTGTGGAGT	TGTTTGAGGC	1100
CCCGTGGAG	TGTGAGGAGG	AGGAGGAGGT	AGAGGAAGAA	AAAGGGAGCT	
TCTGTGCATC	GCCTGAGAGC	AGCAGGGATG	ACTTCCAGGA	GGGAAGGGAG	1200
			C		
GGCATTGTGG	CCCGGCTAAC	AGAGAGCCTG	TTCCTGGACC	TGCTCGGAGA	
			T		
GGAGAATGGG	GGCTTTTGCC	AGCAGGACAT	GGGGGAGTCA	TGCCTTCTTC	1300
			C	T	C
CACCTTCGGG	AAGTACGAGT	GCTCACATGC	CCTGGGATGA	GTTCCCAAGT	
GCAGGGCCCA	AGGAGGCACC	TCCCTGGGGC	AAGGAGCAGC	CTCTCCACCT	1400
GGAGCCAAGT	CCTCCTGCCA	GCCCGACCCA	GAGTCCAGAC	AACCTGACTT	
GCACAGAGAC	GCCCCTCGTC	ATCGCAGGCA	ACCCTGCTTA	CCGCAGCTTC	1500
AGCAACTCCC	TGAGCCAGTC	ACCGTGTCCC	AGAGAGCTGG	GTCCAGACCC	
C					
ACTGCTGGCC	AGACACCTGG	AGGAAGTAGA	ACCCGAGATG	CCCTGTGTCC	1600
CCCAGCTCTC	TGAGCCAACC	ACTGTGCCCC	AACCTGAGCC	AGAAACCTGG	
GAGCAGATCC	TCCGCCGAAA	TGTCCTCCAG	CATGGGGCAG	CTGCAGCCCC	1700
CGTCTCGGCC	CCCACCAGTG	GCTATCAGGA	GTTTGTACAT	GCGGTGGAGC	
T		G	A		
AGGGTGGCAC	CCAGGCCAGT	GCGGTGGTGG	GCTTGGGTCC	CCCAGGAGAG	1800
GCTGGTTACA	AGGCCTTCTC	AAGCCTGCTT	GCCAGCAGTG	CTGTGTCCCC	
AGAGAAATGT	GGGTTTGGGG	CTAGCAGTGG	GGAAGAGGGG	TATAAGCCTT	1900
TCCAAGACCT	CATTCCCTGGC	TGCCCTGGGG	ACCCCTGCCCC	AGTCCCTGTC	
CCCTTGTTCA	CCTTTGGACT	GGACAGGGAG	CCACCTCGCA	GTCCGCAGAG	2000
CTCACATCTC	CCAAGCAGCT	CCCAGAGCA	CCTGGGTCTG	GAGCCGGGGG	
		T			
AAAAGGTAGA	GGACATGCCA	AAGCCCCCAC	TTCCCCAGGA	GCAGGCCACA	2100

FIGURE 2A

GACCCCTTG	TGGACAGCCT	GGGCAGTGGC	ATTGTCTACT	CAGCCCTTAC	
CTGCCACCTG	TGCGGCCACC	TGAAACAGTG	TCATGGCCAG	GAGGATGGTG	2200
GCCAGACCCC	TGTCATGGCC	AGTCCTTGCT	GTGGCTGCTG	CTGTGGAGAC	
AGGTCCTCGC	CCCCTACAAC	CCCCCTGAGG	GCCCCAGACC	CCTCTCCAGG	2300
G					
TGGGGTTCCA	CTGGAGGCCA	GTCTGTGTCC	GGCCTCCCTG	GCACCCTCGG	
GCATCTCAGA	GAAGAGTAAA	TCCTCATCAT	CCTTCCATCC	TGCCCCTGGC	2400
			C		
AATGCTCAGA	GCTCAAGCCA	GACCCCCAAA	ATCGTGAAct	TTGTCTCCGT	
GGGACCCACA	TACATGAGGG	TCTCTT			2476

FIGURE 2B



ISOFORMS OF THE IL4R $\alpha$  PROTEIN

MGWLCSGLLF	PVSCLVLLQV	ASSGNMKVLQ	EPTCVSDYMS	ISTCEWKMNQ	
PTNCSTELRL	LYQLVFLLSE	AHTCIPENNG	GAGCVCHLLM	DDVVSADNYT	100
		V	T		
LDLWAGQQLL	WKGSFKPSEH	VKPRAPGNLT	VHTNVSDTLL	LTWSNPYPDP	
NYLYNHLTYA	VNIWSENDPA	DFRIYNVTYL	EPSLRIAAST	LKSGISYRAR	200
			H		
VRAWAQCYNL	TWSEWSPSTK	WHNSYREPFE	QHLLLGVSVS	CIVILAVCLL	
CYVSITKIKK	EWWDQIPNPA	RSRLVAIIQ	DAQSQWEKR	SRGQEPKCP	300
HWKNCLTKLL	PCFLEHNMKR	DEDPHKAKE	MPFQSGKSA	WCPVEISKTV	
LWPESISVVR	CVELFEAPVE	CEEEEEVEEE	KGSFCASPES	SRDDFQEGRE	400
			A		
GIVARLTESL	FLDLLGEENG	GFCQQDMGES	CLLPPSGSTS	AHMPWDEFPS	
			R		
AGPKEAPPWG	KEQPLHLEPS	PPASPTQSPD	NLTCTETPLV	IAGNPAYRSE	500
SNSLSQSPCP	RELGPDPLLA	RHLEEVEPEM	PCVPQLSEPT	TVPQPEPETW	
		P			
EQILRRNVLQ	HGAAAAPVSA	PTSGYQEFVH	AVEQGGTQAS	AVVGLGPPGE	600
		R I			
AGYKAFSSLL	ASSAVSPEKC	GFGASSGEEG	YKPFQDLIPG	CPGDPAPVPV	
PLFTFGLDRE	PPRSPQSSHL	PSSSPEHLGL	EPGEKVEDMP	KPPLPQEQAT	700
		S			
DPLVDSLGS	IVYSALTCHL	CGHLKQCHGQ	EDGGQTPVMA	SPCCGCCCGD	
RSSPPTTPLR	APDPSPGGVP	LEASLCPASL	APSGISEKSK	SSSSFHPAPG	800
		A			
NAQSSSQTPK	IVNFVSVGPT	YMRVS			825

FIGURE 3

## SEQUENCE LISTING.

&lt;110&gt; Genaissance Pharmaceuticals

Denton, R. Rex

Nandabalan, Krishnan

Chew, Anne

Stephens, J. Claiborne

Duda, Amy

Windemuth, Andreas

<120> Drug Target Isogenes: Polymorphisms in the Interleukin  
4 Receptor Alpha Gene

&lt;130&gt; MWH-0002 PCT IL4R alpha

&lt;140&gt; TBA

&lt;141&gt; 2000-07-12

&lt;150&gt; 60/143,435

&lt;151&gt; 1999-07-13

&lt;160&gt; 413

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 28690

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

```

tgtgagctac tgtgtctggc ctgaataata aaatttataa caatttttca aaaattcacc 60
atgagggtctc actatattcc ctaggctggg ctcaaaccac tggactccaa gtgatccacc 120
ccaccttccc gagtagctgg gactagagat gcacaccatt gcaccaata gagcaatacg 180
tttctgttct ttgtaaatta cctgctctaa ggtatttttg ttatagcagc ctatatggac 240
taagctgact tgtaacgta cttgagactt taaagtgttc cggtcactgt tggagggctc 300
tgtctgtgtt agctcattta atccccacaa cacctcaatc agatggggct attcttagtc 360
ccactttata gataaggaaa ctgaggcatg gaagcacagc ttgctcaagg ttcacatcta 420
gtcagtgaca gagcaggat ttaaacctca ggaaataatc agagaaacat gtgtagaggg 480
ttgtccaagg aaggccacat ccagaagcat ctcccaggac agttgttgtg tagctcacc 540
tctggacttt gtgggtctgg gtgttgtttc atgattatag agagagctct gtgaacgtgg 600
aggacctgtt gtcggcagag acacaaatgg ccaggggcatg gctgggcagc cgcagtggct 660
caggcctgta atcccagcac ttcgagaaga ccagaggggc agatcatgag gtcagaagtt 720
caagaccagc ctggccaaca tggtgaaacc ccgtctctac taaaaataca aaaattagcc 780
aggtgtggtg gtgggcacct gtaatccag ctactcggga ggctgaggca gaagaatcgc 840
ttgaacccgg gaggtggagg ttgcagtgag ctgagattgc accactgcac tccagccttg 900
gagacagagc gagactctgt ctcggaaaaa caaacaacaa agcaaacaaa caaacaacaa 960
aatggccagg gcaggggagg gttgcatatt gaataagatg agctctgctg gaagcacagg 1020

```

tcagcactaa cctgcttccct ctctctctgc aggtgccttg gcatctccca atggggtggc 1080  
 tttgctctgg gctcctgttc cctgtgagct gcctggctct gctgcaggtg gcaagctctg 1140  
 gtaagtcacc acttctcaat cattcatttg ttggctatta atggcgtgcc agggctcctgc 1200  
 agtatgtcac ctggccttat ggagattaca ctgcagtggg aggggacagc caatgacaag 1260  
 tggccctgat taccagtaaa ttctaaagat tgtagaaaag tgatgggagc cgggtgcagt 1320  
 ggctcacacc tgtaatccca gcacttcagg aggccgaggg agggagatcg cttgagccca 1380  
 ggagtccgag gtcagcttgg gcaacatagg gagaccttgt ctctacaaat aataaaatat 1440  
 tagccagggtg tggcagtgca cgctcttagc ccagctact caggaggccg aggtgggagg 1500  
 atcccttgaa ctcaggagggt caaggctgca gtgaactgtg atcgcgccac tccactccag 1560  
 cctgcgtgag aaagtgcagc cctgtcaaaa aaaaagagaa ggtgatgggg aaagaacaca 1620  
 gaacagcata agaggggggtt ggggaagctg ggtggagtgg gggggattgc agttgaaagt 1680  
 aggggaagtca ggggaaggcct cattgagctg acttgaggga agcgggaacc gtgcagatgt 1740  
 ctggggaagg ctcatctctg gcagagaggc cctgcactga gcctggcggg agggttgagc 1800  
 acaggaggga atgtggtgga ggagagtgcg cagcaggagg gagcagtga ggtcagcaag 1860  
 gtgacagagt ggctgaatca aaaaagacct tgcaagtgtt gagcagagga tccatcatc 1920  
 ccattatgtt ccaaaggact cttcaggatg ccgtgtggag aaaggaaag ggtggaagcc 1980  
 aggaggtctg gagggaggtc tggagtggag gagatgagag gctccggatc cctctgggag 2040  
 gtgatttga ggacagattg gaattgaggt gaaagacaga gaaagagaag tggccaggat 2100  
 gactccaaga tttctgacct aaactactgg gaaggacgag gttgtcattt ctgaaatgca 2160  
 gaaggatgcc agaagagaag gtactttggg gaggggcggg aatcaggagt tagttttgga 2220  
 catgagataa gcttggataa tttatttgcct atctaagaca gctccttaac atggttaagcc 2280  
 cttatgcaag ttgtgttcag ctgagatggg cgtggcactg agcatgggag catggaggcg 2340  
 cctgagtggt ctcatgctca ggtggtttag caaactcagt gtacatcctg ccaattccag 2400  
 tctgccatg gccactgaca agctaggagg gcgctgaaag gagaaggacc ccgatgtctc 2460  
 ctccagccca tccatctcct ctctccatt ggccaaacc aaccggaaac taaaggccaa 2520  
 gggtaaccgg tgatgaagac tgtggtatca gcctcctgag cacagagagg gcagaaagg 2580  
 gtggagacaa agagggggcg agatagtggg caaatgggga agtggcactt cccctagctc 2640  
 gagggcagag gcttggtgtg atggaatggc actccttaaa ctgctacata ttttcccttt 2700  
 aatttgcca agaacaagt gtcaagtgtg tgtgagataa aggtgcactt ggttcgttct 2760  
 tgtctaattg ccccgccacc catgggtatt tcttcagctt ccacagtcac cccgacacta 2820  
 gctgggaagc tccagcagcc ctggtcctgg cccagctct gtggcgctg gccctcaact 2880  
 ttgctgcac tgtgctttg tgctattccc cttggtcctg tttgggtgca agtccccctc 2940  
 acgcattgag ttcctgggac gctcaggctg ctctgtgtc tcccaggga acatgaagg 3000  
 cttgcaggag cccacctgcg tctccgacta catgagcat tctacttgcg agtggagat 3060  
 gaatggtccc accaattgca gcaccgagct ccgctgttg taccagctgg ttttctgct 3120  
 ctccagtaa gcctgcgctg gagctggagg tttggggagg ttgtgcccga agggtttgcc 3180  
 ccaagagtga gctgggtcca ggtggtgcg tggagtgcag gatgctgagt atggtttgct 3240  
 gctgtttata tgggtgttag ggggaggtcc catctccagg gacatgttat gtaagataca 3300  
 gtggagcgca tgggtggagt gttggtccac gtggcacatg gatacggctg gaatactgga 3360  
 ctagaccagc agttctcaca ctttttggtc tcaggaccct ttttcacact taaaaatgag 3420  
 tgaggaccca aagggtttg gtgtaggtaa cacatcatc tatgtttacc taattagaac 3480  
 ttgcaatgaa gaaatggtgt aatttttaaa aaattaaac aattaaaaat ttttttctt 3540  
 actgaaatgg aggtctcact gtgttgccca ggctgctct aaactcctgg gctccagtga 3600  
 tctcctgcc tccgctccc aaagtgtgg gattacaagc gtgagccgct gtatccggcc 3660  
 caaatggag aaattttaag tcccaacaac atgcaagccc gcattcaaca aatcttcaga 3720  
 tcaattacat gatcacagg catgtagcct ctagaaaatt cactgtacg ccagtgcag 3780  
 agagtgaata ggcaataac gtccctgtat tatgatgaaa agagttttac ctggtgggcc 3840  
 cagaccacac tttgagaacc actggactag acccttgatt gaggagtac gtgttgagag 3900

tggagtcctc tgtgatggtg gatggaccag gacacatggc ataggagtca ggtggttccc 3960  
 tgggctactc catggtgcac aggatgcttc gttacactgg tgcccaggac ataatacacgt 4020  
 acacaagaca cacagttacg gggcagactg gggatatacg gcacaccagc atgcagcggt 4080  
 caccagtaaa ggtggtattc catgattatt ctaaggtaga tgggctgtgc tttgtttcca 4140  
 ttggcttagt ccagggattg gcaaactatg gcccgtagc caaatccggc ccactgcttg 4200  
 tttttgtaaa taaagtttta ttggaacaca ctggctgctg tagttgtaac agaaactgca 4260  
 tggccctcct ttatgttttt tgtttgtttg tttgtttgtt tgttttcttt gagacagagt 4320  
 ttcgctcttg ttgccaggc tggagtgcag tggcacaatc tcggctcact gcaacctctg 4380  
 cctcccggtt tcaagcgatt ctctgtctc agcctccga gtagttggga ttaatggtgc 4440  
 ctgccaccac acccggttaa ttttcgtat ttttagtaga gaccggtttt catcatgttg 4500  
 gccaaacttg tctcgaactc ctgaactcag gtgatccacc cgcctcagcg tcccaaagt 4560  
 ctgggattac aggcattgagc cactgagccc ggccctctcc tttatcttaa ttgaaataat 4620  
 tcagaaatgg aaagtcaaat actgcatgtt ctacttata agtaagagt aaataatgtg 4680  
 tacacatggg cattattcca tgtaccatgg aataacagac attgaagact tgggagggtg 4740  
 ggagaggggt gaaggaagag aagttactta atggcatag tgtacaccat ttgggtgacg 4800  
 gaccacccag aacccagac ttcaccacta ggcagcatat ccagtggaga cagatctgag 4860  
 gcttgccatc aaaattgcac ttgtaaggcc gggcactgtg gtggtcgcg gctgtaatcc 4920  
 cagcccttg ggaggccgag gtgggcagat cacttgaggt caggagtctg agaccggcct 4980  
 ggccaacatg gtgaagctcc atctctacta aaaatacaac aattaactgg gtgtagtggc 5040  
 gcacacctgt aatcccgct actaggaggg ctgaggcggg agaattgctt gagcccagga 5100  
 ggtggagggt gcagtggacc gagatcacat cactgtactc tagcctgggt gacagtgaga 5160  
 ctttgtctca ggaaaaaaaa acaaaaaaaa aaaacaaaaa actcgtaccc cctaaattta 5220  
 tacaataaac caaaaaaaaa aaaaaaaaaa gaaatttgtt ggcccttgaa gtccaaaata 5280  
 ttaactatct ggccgtgtac agaaaaagt tgcagacccc tggcctagcc cgtgagatgt 5340  
 ggggtggtg ttaaggtgga acattggaat tatcttacga tggccaaact gtgcgatgca 5400  
 gagcttatgt tgttctaaat taattagtgc caccggttct tccctttcat gggctttcag 5460  
 gaacaagcta agtcccagga ccaggccgg cagctaggca ggtgtgagga gcatccttg 5520  
 tgcattgtgt aagaggctgt ggccagcaag agaggcaacc ctagtgggt gccccagcac 5580  
 accctggccg ctcccaagcc ccagatctg tctcacatc cgtgatcggg aagctggaag 5640  
 agtctgatgc ggttcttgga ggcattgccc ggacacagct gtggggccca gccagcctac 5700  
 aggtgaccag cctaaccag ccctgtgtc tgcagagccc acacgtgtat ccctgagaac 5760  
 aacggaggcg cgggtgctg gtgccacctg ctcatggatg acgtggtcag tgcggataac 5820  
 tataactgg acctgtgggc tgggcagcag ctgctgtgga agggctcctt caagcccagc 5880  
 gagcatggtg agcaggcgcg agtgcggcag ggggtgctgg gtgtgttccc acagctgcct 5940  
 gggctgaggg tgggtggggc aggggaggag gtgggtcat agcaacagca ggaggaagcc 6000  
 gcctgtattt tcccaaatct gatgggattc ctgccctgc ctgggcctca gtcctccac 6060  
 ctttgaaacg gagctggtcg cagtagacca ccaagcccc ttcagcccag ctgtttccac 6120  
 ccctgaactt aagtgccag gaaggcgtat tgagatgagg tgtgcttgct ggaaggcatg 6180  
 cctgtgctg attgaaaacc gaactgggaa cattccttcc attctgtgtc cactggtcag 6240  
 ctgctgcggc ttggatggt cttgaccgtg gaaggctgac cttcttctgg taccggagt 6300  
 ccctgcagga atcccccttg agcttgctgg gctgtggtga caggagtta aaacatgct 6360  
 tgtattccag tgatgcatga tatgacatgc atcacaggaa taaaaacctg aggtctcatg 6420  
 gatatgattg cttcaaagga gaccaagttt taaaacagat gaatcaaat aaagaaaaat 6480  
 actcagtaaa tcatcataaa gtacagagat gtggccaaag gtgtgaagga tgcagctgta 6540  
 aaagctgaag tttagggccg ggtgtggtg ttcattgccta taatcccagc actttgggag 6600  
 gccgagccca gcggtacacc ggaggtcagg agttcgagac cagcctggac aacatggtaa 6660  
 aacccgtct ctactaaaaa tacaataaat tagtctggca tgggtgcagg cgcctgtaat 6720  
 cccagctact tgggaggctg aggtaggaga atggcttgaa cccaggagaa ggaggttgca 6780

gtgagcttag atcatgctac tgcctccag cctgggagac agagtgagat tacgtctcaa 6840  
 aaaaataaaa ataaataaaa ataaaaagat tttttaaaag gctgaagttt gggttacttt 6900  
 ggctcataca ctttgccttc actgtagaaa ggtgggttagt aaagaccagg cgcggtggct 6960  
 catgcctgga atcccagcac tttgggagcc cagcgagagg agatcacttg agccctgggc 7020  
 tattgaggct gcagtgaact gggattgtgc cactgcactc cagcctgggc aacagagtgg 7080  
 gaccctgtct caaaaaagaa gaaaaaaagg gtaattaata aacactaaag ttctatgtag 7140  
 aatttttagca acattattgt tattataatc ttctttgcta tggctctgaa tctgtgtggt 7200  
 gctccagaag tatgctatgg aggttttgtc gaccaaaaat ctgggtgggtg gctgtgtggtt 7260  
 gtaggccggg gctgggctgg gtgatggggg agtactgca tagatcctca catagaggcc 7320  
 gcttctcccc cagtgaacc caggggcccca ggaacctga cagttcacac caatgtctcc 7380  
 gacactctgc tgctgacctg gagcaacctg tatccccctg acaattacct gtataatcat 7440  
 ctcacctatg cagtcaacat ttggagtga aacgacctgg cagatgtgag tgggcatgct 7500  
 ttgacgtttt tctgtgacct ctgggggaaca ggggtgggtg ccagcagagg cccagtccct 7560  
 ggagccagga gcctgggagg caagccctgg ggctggatag caaatccag gagctagaga 7620  
 cctggcctct cacctggctc tgccctaggc aagtcctttt gcttctggc cccccaccc 7680  
 tcacatcaga gaaggggagt tatctctgca tgccgctcct cctctgtaa ggtagggtg 7740  
 tgggccacat ctgtgtttcc cagtttgggg gacacaagt atcgtagggt gcacattgac 7800  
 agctcacttg aataacctt ttattgaaga gaataatact gactcaagag acagtgacct 7860  
 gtgtcagttc ccttttgagg ccaacgggtt aaggaggaag tccccataga gctgactcgt 7920  
 ttactaatc ctcttaatga agagagcaga ggccacaccc caggcttaga ctttcccaag 7980  
 aaaacaagat cagtttgttg gttgttcccc atggaagctg gtcctgacat tcccttcaca 8040  
 gtagtgttg tggagtttt gttgttgtt gttttgagac agagtctcac tctgtcacc 8100  
 aggggtgaac acagtggcgt gatcttggct cactgcaacc tccgcctcct gggttctagc 8160  
 gattctcctg cctcagcctc ctgagcagcc gggactacag gcacctgcca ccgtgccag 8220  
 ctaatttttg tatatttagt agagatgggg ttctactgag ttggccaggc tggctctaaa 8280  
 ctctgacct cagatgatcc actgccttg gcctcccaaa gtgctgggat tacaggtgtg 8340  
 agccaccgca cctggccagt ggagtccctt cttaaagtaca tgtattgaca tctttaaaaa 8400  
 gggcgagagg atttacagga aactatcagg tcagtaaatg caggggcccgt ccacagtggg 8460  
 tggctgagtc cccctatttt tctgctggtg tgcaaggagg tcatttctct ccacccatgt 8520  
 ttccccaccc tgaatccacc ttctcaccat tcccatgga gggacaatct ctggacatat 8580  
 gggacctggg gtcccacagg gctgcaatcc aatgcctgct gtgccactcg ccagctgtgt 8640  
 gatgttgggc atatcccata acctctttgt gcctcagttt cctcatctgt aacacaggag 8700  
 tgacaagagc acccgccac agggctatga cagtacaagg tgtgtgatac agatgagctc 8760  
 cctgttttg cccacatgtg tcttaaaagc catgtgccct ttctcttgag tgccccaggc 8820  
 cacagagatc cccatctgcc cgtgtccca cactgggtc tgtcatttgt tcttgaggt 8880  
 ttgtgagggc cggctctgtg catcccagg gcccaggctg ggctgggtg gctctcaggg 8940  
 agcaggcacc cgccacctta agtcccatg ctggtgtctg tctactgttc ctctcaatct 9000  
 ggccaagcca ggggtgtcga ttatatctc tcaggctggg ttccccctt ggcactgggc 9060  
 cagggtatggg gaaagagcag gaatggggca gttggctcac acagcagagg ctcagaaagc 9120  
 ggggggcatg gggggaagga gtgcacagat gctagagagt ggggcaagtt ttgtttggtc 9180  
 aataaatctc cttctcatgc cccaggcctg tgcaagacct acagagagtc ccaaggatgg 9240  
 gctgggggga agagaaaggt accaccttca ggtccaaag atatgttatt taatattttc 9300  
 atatttctag atctgccttc aggcattggt ggatccagct tctaggaacc tgtccagctc 9360  
 tgcgcccctg tttattctgt attgcttcg ttttaggca ggctcttccc tcatgtagt 9420  
 gcagatatgc ctactagtgt ctccaggcct acatcccaaa gccacagtgg gaaaagggtt 9480  
 tttttcttg acggttctaa taagagctct aaggctgctg ctcagtggcc tggcttcgat 9540  
 gctgtgccag cctctgaacc aatcactggc tgtgggtgga gagagggtgc tgggtggagg 9600  
 ccctgcttgt ccaggaggga gtcacatacc tgcccttagg gctgcagggt ggctcagctc 9660

catccaaacc agatgaactg aaaataaggc aggagtggct tccccagggg aaactgggga 9720  
agaggaagca ggactgtgct ggctaaaatg ccagccaggt ttaagacgtg gcaccagatg 9780  
ccagtcattg gattggattg gtcagcatgc ctgggctatg gcttaggggt atgttggtgc 9840  
tcagggatgc cacaggcctc cagataccag gtctgaggca gaagaatgaa gtccagcttc 9900  
tcttgtgggt ggaacagtgg caactgagat accccatctc tcccttccca agaacagagc 9960  
tgaacataaa gaatttagtg attggccaga gcttggccac atgctccct ctgatgaatg 10020  
ataggccagg tgatgggatt ggcacaattg gcttagacta atgaggggtg gccctggagt 10080  
tgcaggcagt ggagtctgt cctaagcagt gggcacctaa acccgatggc ataaaagctg 10140  
ggcgggtgtc cacctgcctc tgccacagca ctataggcac caactgtggc tcatactgag 10200  
tgggataaat tccagaaaga aacattagga acttactata gaattttggg gctagagcta 10260  
ctcattcatt ccctagata atttctaggc aaggttccat agtggagggg gagttttggc 10320  
ttgggcattg aaggatgcat aggagttttc tagatgggga aagaagggaa cggtagacca 10380  
ggcagagggg actgcatgat aaaaggttta tgggtgtgaa aattcatgga atgtttgagg 10440  
attatggggt tgggggatgt gggaaatgt gttagcgtaa agcaccaaac aaagccaaaa 10500  
gtttagttag agccctgaat gcctgcctca taatggtttc catattttat atgcctacta 10560  
tgtgccaggc acattgctca ggtcacaca gctggaaatg gcagggctga gttttgttg 10620  
ttgttgttgt tgttgagaca gagtctcact ctatcaccca ggctggaatg caggggcgtg 10680  
atcatggctc actgcatcct tgacttcctg ggatcagggt attctccac ctctgcctcc 10740  
caggtagctg ggactacagg cacaggccac cagccaggc taattttttg tatttttagt 10800  
agcgacaggg tctcgccatg ttgtccgggc tggctggat ctctggctt caagtgatec 10860  
ccctggctca gccctccaaag gtgctgggat tacaggett agccaccgca tccagccag 10920  
atctgagatt tgcacccagt atttgaactc ccaagcctgt gctcttttct ctcccatgga 10980  
catttctctc agagatggtc tcccaaacac ctgtccttct tgttaaaaaa cagacaaacc 11040  
gcaagtagtt ctttgggaagc tcagatttct cttttgttct ttagtaaaac atttcccagt 11100  
tcccagctcc cttccagggt gtaagatttc ttccgtaact tacatctagc tgttgcttct 11160  
tgtttgctca tgtttagaaa gaaagacaaa agagagttag aattttctct cccttcccca 11220  
gtctccccac aactcacacc ccaccctcag ctccctctgt aataggaaaa tctctgaact 11280  
ctctgtagtt gctccagcaa tcttttgga ctttgcctct ttcttgtaga aaaacctccc 11340  
cttggctcac tttgcaccag gtttcccaa atgtgcttcc aaccacaagc agaaatggag 11400  
ctgccagtaa ccaggaagaa actgccgggg gctgaggaag aggagaggga ggtgcatagc 11460  
cctggatctc gcagggagag ggtgacagg atgagaactc aggttgctca cttgccatca 11520  
gggtcagtea tgaatatagc gttcatgtat cactttttaa agcttttttg gagggtaaaa 11580  
gtaatagtta cacaaaataa aaatacaaat ggtacaaaag gacttagaat ggaaacatgt 11640  
ttctctcccg actccagcct cctgttttct tcccagaga ctgaccactg ctgtctgtct 11700  
cttgccagaa gggaaaggga ggcaaggtaa gggcaggcag agggcatgtg catcctttag 11760  
agagagctta tgtctataca agcaaatgtg tgtgttcagt catcgtctgc ttagttttct 11820  
attgctgcat aataatggta ctaccagctt cacagcttta aacaacaccc atttattatc 11880  
tcatagtctc tgtggttggg agtctggaca tagcttagcc aggttctctg ctttagagtc 11940  
tcgtgaggct ataataaggt tgtgggatgg ggtgcagtt tcatctgagg ctcaattggg 12000  
gaagggtcac ttctaagctc atacaatatt ggtgacattc agtccctggc aggctgttga 12060  
actgagagcc tcagtttcgt gctggctgtt ggttgtagtt aaccctgaat tccttcccat 12120  
gtgccctttg caaagccatc aaggcagaga gacttgctta gcaagtagga tattacagtc 12180  
ttctgttaata taatcacatc catgaaatcc tctatatatc ccatcacctt taccatattc 12240  
tgtgggttag aaacaagtag caggtcctgc ccacactoga gaagaccaga tgacacaaag 12300  
atgtgattca aagtggggat catcggggcc atcttaggtt tgtctgcagt gatcactgtg 12360  
ccatctctct ctctctcttt tttttttttt tttttccgag acgaagtcgt cactctgtca 12420  
cccaggctgg agtgcagtgg catgatctca gcttaccaca atctctgcct cccaggttca 12480  
aatgattctt ctgcctcagc ctctgagta gctgggatta caggtgcccg ccaccacacc 12540

cagctaattt ttgtatTTTT agtagagaca gagtttcacc atgttggcca ggctgggtctt 12600  
 gaactcctca cctcaagtga tccacccact tcggcctccc aaagtgtctgg gattacaggc 12660  
 atgagccacc atgccagcc ccatctctct ttaaaaaaca aacaaacaaa caaaaaacat 12720  
 aaaaagaagc agagaacaca tacacatctg catcttccct tgtttactta acaatagatc 12780  
 ttggaagtca cttctcagta gaggttaggt tgggcagagc attggattct aggccagtga 12840  
 gtttggactt gaccatggag aacttaggaa gcccatgaag gacagagaga gatgcctcga 12900  
 ccctgccagt cctttagaaa gatcacccag tgctttttgt ataccaaacc ctatttgaaa 12960  
 tacttacgta tattaaccca ttctcttctc accacaaccc tgcgggaagg gagataggca 13020  
 cttttattat cttcattttg cagatgagga cattgagggtc cagagagggtt atgtcactta 13080  
 ctttaaggta cacagccagg aagtggtagt agggactctt acccttgttt tacagatgag 13140  
 attgaattat ctacagaaaa ctcagaaagg ttaaacaaact tgcctaagta acatacagct 13200  
 aattagtctg ggagcctgac gcatgttgct ctgacctggt cacagttaca gaggtggcaa 13260  
 gcaatggcct gaaaggagc aacaacaaaa taccagggtt ggtggctctt aaacatgggtg 13320  
 ggttcagcta acgacagcaa ccagggtggg cactgggtgcc cctcgccccc ggctgggtggc 13380  
 ctaacatctc cttttctctt accagttcag aatctataac gtgacctacc tagaacctc 13440  
 cctccgcctc gcagccagca cctgaagtc tgggatttcc tacagggcac ggtgagggc 13500  
 ctgggtctcag tgctataaca ccacctggag tgagtggagc cccagcacca agtggcacia 13560  
 ctgtgagtat caagaggcct aagcaatggt aatctccact ctccattctt ccctgtggc 13620  
 cagacacttc ccctggctga gtctctgggc ttttatatca taggatgcct ctaatggcaa 13680  
 tcttgccatt agatacacct gctgtggtgt atctgccagg taggcaggct aggtctcagt 13740  
 aacacacaag cccacaattt ccatggctta aactataggt aatataattt ttgctcatgt 13800  
 aacaagctaa cgtgaatggt gctggtttgt aggtggtttt cctccctgta gaaatctggg 13860  
 gagtgaggtt ctttccatct tgtggtgcca tcattctcca ggacaaagat tcttacctac 13920  
 ttttgtgtcc tggttttctt tggcagcctg gtgaagccta tggacctcat ttcagaatat 13980  
 ttttaaatat ataaaatccc agcctgggca atatagttaa acccccatct gtacaaaaat 14040  
 tagccaggca tggtggcatg cacctgtagt cccagggtact gggaaaggct aggtgggagg 14100  
 atcacttgag cccaggaggt tgaggctgca gtgagccgtg atcgaccac tttactccca 14160  
 cctgggtgac agagcaagag cccatctcta aaaataaata aatacaatga aataaaataa 14220  
 aataaataga actacagagg aaactaattg tattgaaatg cagtataaa acatttaaac 14280  
 acatttttaa tctagagata tatgtgcttc tttattaaga tctataaata ataagttcta 14340  
 ggggtagctc gcataaatat tgtaatttca aagtagataa gcataaataa tactttatga 14400  
 tactgaaatt gtgatgtgat atgagaatag ctgtgagttt tgttttgctg gggacaggat 14460  
 cactgatgct gtcattactg ggtctcttc cctccattct ttttttaaaa ttgtatttta 14520  
 ttttattttt aaaattttta aataaataga gacagggtat cactatgttg cccaggctgc 14580  
 ttttgacctc ctgggtctca gtgatcttc catcttggt tcccaaagt ctgggattac 14640  
 aagtgggagc cagtgttctt gggcccttc tccattctta atggaaggag atgctagggtg 14700  
 tgagagggtta gggaaagtaa agatgtaatt tctttcccat ccaagttctc agaccctga 14760  
 attctacctg cagccatggt ggtccatcaa ccccaagtga agaaccctg ctctaggggc 14820  
 ccaccattgt ctgtatccag ccagcagaag aggcgtgatt atggagatca catctgcttc 14880  
 ttgaaagcag acagcccga agtgggccgc atcacttct ctcaaattct attggtgaaa 14940  
 atggtcacat gactacacat agccacaaag gaggtggga actttctcac ttggaacctt 15000  
 catcccagaa acaactcttt tcagtgggt atcccacagg tctttcgag tagaaatatt 15060  
 gattatctca cataaatga agtcttaca atggacctac tgggtttgt acagcagcca 15120  
 agtgatatct ctccctctt gctgtcttc cttctgcgt cttcacatg gtggcattgt 15180  
 atccttagac ttgccacca tggcctcagg ttggccgttg cacactgtct tacataaagc 15240  
 aggaaggaaa ggaaggctg ctacgagaga gtgtacctg tgcattctt ttttaatcag 15300  
 gaagcaaaca tctttctaga agcttcccta gcaaaattcc cttacatct cattggccaa 15360  
 gactgttaca tgttacatgg ttactgttat tacttgcctt ttgcaaggaa gactgggaac 15420

tcaaattgcct ggaaaaagga acaggataat cgtgattggc tcaagcctta ggggtgggcat 15480  
 ggctccctga caaggggagag aggaaaaagc tgttgagtga agaagactgc ttcagtttcc 15540  
 ccatctgtat aatgggagga gtaagggctg tcgtgaaaac tcaatgaaag aagattcttc 15600  
 aacgtggtag gtgcagtggc agctggcagt accctgaccc tgccaccgca cagccctctc 15660  
 agcattgctc atcctgcact gtggatatca gttgagccac gtgtctcctg ccctgggctg 15720  
 tgagctccat aggcagggtc tccatggctg tatctccaga acccagcaca gaaccagggtg 15780  
 cttgggaaag ttttgaattg attctcatct gccattggca tggggaaggg aactagcttg 15840  
 tatgaaacag ataacaatgt atgggaccct cattcattat ttcagcaaatt atttgcctgag 15900  
 ttctctctac atggctagcc ctgtgctaga cactggggaa tcggcgatga acaaagcaga 15960  
 tagaaatccc cactcttggt gagctgacat tctggaggga gagacaaaaa gcaaacatat 16020  
 aaagaaagaa agaaatcaca tggatctgga tgacagttag tgctgggaag aaaataaaag 16080  
 cagaggaagg ggtaggagcg atgggcaggg ggcaacggta gggaggggtg cggggaaaac 16140  
 tttttggaga atgtgacgat gaaagtgaac aaggagaagt caaccgtgtt gagatgatgg 16200  
 cagctaataga tgtggacagg ccactctgtt ctgagtgcac tatctattga ttcactatgt 16260  
 catctctgca acagccctgc acgatcaatt ctgtcattaa ccccatagta cagatgagga 16320  
 tgcggaggca cagagaagat aagggacttg tcctgtgtca cacagcaagg agccatccgg 16380  
 ctcttaagtt ggtgcatttg acttctgtgc ttccggaaag aaagagcagc aagttaaaga 16440  
 tctggagggt gcactgagct ttggaggagc agggggcaat gaggtggcgc gtgtgacgag 16500  
 gactcaatgt gcaagaggga gagtgggtgg gagatgaggt ggagggtggg tcggcggtca 16560  
 gatcgtggag ggtctcggac gagggctctg accctgggtc tccagtcctg ggaagtggag 16620  
 cccaggtctg accatggctg acctcagctc atggtctccc ctcccacttc cagcctacag 16680  
 ggagcccttc gagcagcacc tctgtctggg cgtcagcgtt tctgtcattg tcactctggc 16740  
 cgtctgcctg ttgtgctatg tcagcatcac caagttagtc ctgggcccag tgctgccgag 16800  
 cagtcctctt ggagtgcagg gtggcaggga cttgcccctc tagtctgccc ctttgagtc 16860  
 ctctcagtcataataacgta tttactgagc agctactaca caccttgaga gtagagctga 16920  
 gaacatatcg acaaggaccc cacttttttc tttttttctt tttttttttt ttttgagacg 16980  
 gagtctcact ctgtcaccca ggctggagta tagtggcaca atcttgccca acagtaacct 17040  
 ccgctcccgc ggttcaagca attcttctgc ctcagcctcc agagttagctg ggattacagg 17100  
 cgcattgccac tatgcccggc taattttttg tatttttggg agagatgggg tttcaccatg 17160  
 ttggctcaggc tggctctgaa ctctgacct catgatctgc ctgcctcagc ctcccaaagt 17220  
 gctgggatta caggtgtgag ccactgcacc caaccaggac tccacatttc taaaaccggc 17280  
 atcctaactgg ggagactgaa aatacatatc aatcacaaac aggtggtttt ccatagtgac 17340  
 ccactctctg aatgcactag accagggtgg aggccagaga tcttctgggg tgctttttgc 17400  
 aagggggacc aggataaggc tctccaagga gggaaaattt gagggggggc ctgactgggg 17460  
 agaattgagct ggccagggat aagcaagatg gagtcatccc acatcccctt acaacactgg 17520  
 gtgcctgggc aactgggggc atttgggggc atgtggtagg agccagagga atttgcgacg 17580  
 attgccctga tggagtccag agacctgggt ttgaatcctg gccttgagac ttggtagctg 17640  
 gcggccgaca agttgctgaa acccctgagc ctgggggttc tgctttgcag agtgacagtg 17700  
 atggtgagaa catatttcat cagccagaag aggccaaatc acagtaaagg ctgagggagg 17760  
 agatgagtgg cgagtggctg ggagggtggg gaaggagcct cgtttccaga gagctcttgc 17820  
 cagcccttgg aatcatggtg tctcagagcc tcagtcctcc catctctgaa atgggactag 17880  
 caagctcaac ctactaagt caggattaga ggtggctaag gattattaac atgattgatg 17940  
 aaagtgccca ctcttgccc agcacacact aggtaggcag ggaatgcaa tccccctcca 18000  
 tatcttgta ctgatgcctc cgagcaacct tggactgac gccttgctct gagcctcagt 18060  
 ttccccatca cctgtacctc tccccactcc ccatcactat atcccagcat gccagcctct 18120  
 ttgctgttct ttgtctttgg tttctgttt tgttctgttt ttttagacagg gtctcactct 18180  
 gttagccagg ctgaagtga gtggcgcggt tacggctcac tgcagcctcc aattcctggg 18240  
 ctaaagagat cctcccattt caacttccag agcagctggg acaacaggcg cttgccacca 18300



cacctggcta attttcttat ttttaatttaa ttttatttta ttttttggga cagagtggag 18360  
tctcaaaaac caagctagag tgcagtgggtg cgatctcgac tcaactgcaat ctctgcctcc 18420  
cgggttcaag cgattctcct gccttagcct cccgactagc tgggattaca ggcgtgtgcc 18480  
acgacaccca gctaattttt gtatttttag tagagatggg gtttcacat gttggccagg 18540  
atggtcttga actcctgacc tcaagtgatc caccacctc gttctccaa ggtgctgggt 18600  
acaggcatga gccactgtgc ctggccaatt ttcttacatt ttgtagagac tggctgtcac 18660  
ttatgtagcc caggetgatc ttgaacttct acccctttat ctttattcat ggcacttatt 18720  
accatgaatg aatgacctca tataagcatt tctttcgttt tttttttttt ttctttgaga 18780  
tggagtctca tgttgtcccc caggetggag tgcagtggcg cgatctcagc tcaactgcaac 18840  
ctccgccttc cgggttcaag cgattctcct gcctcagcct cctgagtagc tgggattgca 18900  
ggcgcctgcc accatgcctg gctaagtttt gcatttttag tagagacggg gtttcacat 18960  
attggccagg ctggtctcga acttctgacc tcaggtgata cacctgcctt ggcctccaa 19020  
agtgtggga ttacaggcgt gagccgccat gcctggcctc atataagcat ttctgtctcc 19080  
atttatcatc catctttccc tcttgaaggc cagtttcacc aaggcaggca tctttgtctc 19140  
gttcaactgt gtagcctcag ggccaggcac agtgagtcaa acatagaagg tgctcaataa 19200  
atatgtgttt atttattgaa accatgggca gaggctaatt cagaagcggg ctgaggacct 19260  
tacctcccag tgatgatgca ccatggcccc aggaggcca ggaagagaga agggttgtgt 19320  
ttctccgtag gtccccagc ttcccaggcc atcccaggcc attccctggg catttgccct 19380  
cagctgctct gaaaaaggga ttgttgaggg gaacctagaa tctctctctc gcagtttgag 19440  
tctttcctaa tcccctgggg tctcattccc actgaggaca taggtggcct cctcaggaa 19500  
tctgtgctgg gtaacagaat gcgggagtg gaacctggct ctgccacctc ccagctgtca 19560  
ctccacctcc ttgggcctca ctctcctcat ctgtagaata gggtagcaa tagaatccat 19620  
gtcaccaggg tagaatgatg agtcagtggt ttgacctcca gaaactaatc agcctgatct 19680  
ctgatgcaa ataagtattg gtgataacga ccacttttat gggaggagcg ttcacctgtc 19740  
aataattcag agatcaacac cttttccttt tgtttttcag gattaagaaa gaatggtggg 19800  
atcagattcc caaccagcc cgcagccgcc tcgtggctat aataatccag gatgctcagg 19860  
taggagtagg cgtggatgag gacatgtggg actgtgtaca tgaagaagtg tggttcagaa 19920  
cacctgggct gttaaggacc ttactggct tctggaatgg caaatagaca gtcaggaggg 19980  
ttgcagggga gacagaggca gaagccgaat gaggtcatta gcagaccaga ggctttcccg 20040  
cccttcccct tggcaatccc agcctggggg gggcttctct ggggttggtt tctgttttt 20100  
ttccctcccc ttgggagaat gaccttggg tcatcatcac tgtgtcatc cctggggagg 20160  
tgccagtacc agggctagag gccagaagga gtggaggaag gagagggtga caggctttct 20220  
gtgtcttctt cttaagcata ggaaactgcc ccgaagcac tagcaaatcc ctccgggtt 20280  
ctcattggcc tgaatgtat cccaccctca agccagggt ggagtcagct tcccgaaggc 20340  
gatggtcctg tgggtgagtg ggtggggtt gcctgagcaa gatgagagtt ctctaggtag 20400  
gagaaagggg gattataggt cctgtctaga agagaaggtc tgagggtcct tgcttttcca 20460  
gggactctgg aatctagtgt tggctttgaa tctgactct gccactcact ggcagtgtgg 20520  
acttgagcaa gttgcttaat tctctgagcc tcagtttctt cttgtgggtt ataacagtgt 20580  
ttacctggta ggacagatat tggaaattat tgagacaata catataaagt gcatattcca 20640  
gcctcttgca aataccaagt gccatttatg tatcagttag tgtttgctgt gtaacaaatg 20700  
accccgaaat gtagagggtt acaacaactt tatttagctt atgcttctgc aggctggcat 20760  
ttggggctgg gctcagcagt gaggtggcg ggggagctg ggctgggctg ggctgggcag 20820  
atctgaattg agctgaccog tcccgtagc ctccctccgt gtctgacagt tggctttttt 20880  
tttttttttc tttttctgag acggagtttt gctcttattg ccaggagtg caatggcatg 20940  
atcttggtc actgcaacct ctgctcctg ggttcaagca attttcttg ctcagcctcc 21000  
caagtagctg ggattacagg catgtgccac cagccaggc taattttgta ttttaatatg 21060  
agatgggggt tcttcatggt ggtcaggctg gtctggaact cctaatatca gatgatccac 21120  
ccacctcagc ctcccaaagt gctgggatta caggcgtgag ccactgcacc cagcctagtt 21180

ggctgacttt tacctgggac agtgcagggt cctgagccat gtgcctctca ctctccagca 21240  
 ggccggccca ggcttggtta cagagtggct cagttttcaa gggtggaag tcccaaggct 21300  
 tcttgaggcc tagggcgagc actggcatga tatcacttcc atcacattct atgggcccga 21360  
 gcaagtccca gggccagtgt agattcaagg gatgggagga gattcagagc actcctctgt 21420  
 ggccactttt gccatcgacc acagtccctg taaatattag gacaatgtaa ttaattccca 21480  
 ggaatctgag gctcagaaag cgtaagtgtg ctgttggtgact tctgatctgt gtgatgtcga 21540  
 ggcttgatcc ccttctctgag cattgccgta ctccaggccg ggctgcaagg ccactctgct 21600  
 ctttcatttg ctgtctctgt atttttaggg tcacagtggg agaagcggtc ccgaggccag 21660  
 gaaccagcca agtgcgcgta tgtatctgaa cttaggtcac agcctgcag cattgggaag 21720  
 gtgatagaat tggagaggca agcccctagc tccatgtctg ccttctcttc cctgcattcg 21780  
 gtaattgccc tgtgacatta gccttcaagg gacggcagga ggagggtgt tctggaaacg 21840  
 tggactgctg gccaaagccc ctgagtttca ctggtgtgtc aggtacatgg tgataccct 21900  
 tgggagtgtg gttatagtta acaaccagag cagccgtgcc tgttgtaaa atcttgacct 21960  
 aattgtatac ttgtcggcaa atagccacta tcctgaacac tcccctcctt ttttttaata 22020  
 tacaggatct cactctgttg cccaggctgg tgtgcagtgg tgcgatcata gctcactgca 22080  
 ccttcaaact cctgagctca agtgatcctc ccatcttagc ctcccagta gctgatacta 22140  
 cagatgtgca ttaccacgcc tggctatttt aaaagggttt tgctgtaat tccagctact 22200  
 caggaggctg aggcattgaga atcacttgaa cccgggaggg agagggtgca gtgagcgag 22260  
 attgtgccac tgcactccag cctgggcgac agagtggagc tcttgtctca aaaaaataa 22320  
 taccaaaaaa agtttttgta aagacaagct ctgcgtgtgt tgccccgcca ctgtggcctc 22380  
 cttagcttct tcctggggc ctgctggacc ttccatact ccagaaacta aagggggtcc 22440  
 aggacctgc ttcaacccta ggatcccga tcttttttt ttttttttt ttttgagcg 22500  
 aggtctcttc tgtgtccctc aggtggagt gcagtgatc actgcagcct caaactcgtg 22560  
 ggctcaagtg attctctagc ctacagcctc taagtagctg ggactacagt catacaccia 22620  
 catgccagc taattttcct tttttttaat tctttagag atgtttgaga cggcttgggc 22680  
 tctgttgccc aggtgttct caaactcctg agctcaagcg atcctccctc ctacagcctc 22740  
 taaagtgtg ggattacagg cgtgagccac cgcacccggc ttccatatcc tttctaattg 22800  
 gtcattggctt gggataatgg tgttgctttt aattatcatc atccataaag actttttctt 22860  
 actcaacaga tctgagcttg tatttgggtg ccaggacatg tgcgtgggtc ccgaaatccc 22920  
 aaagacacag accctacct cagggtttc tcattctagc aacatagact gatcaattac 22980  
 tgattataac gttagaaggc atgtctgaag tagacagcca tcaggacatg gtgatttcag 23040  
 gctgggcttt gaagaatgaa taggagtttt tcaagtgtcg aaactgaacc ctgaccaacc 23100  
 tttgcttttg cagacactgg aagaattgtc ttaccaagct cttgccctgt tttctggagc 23160  
 acaacatgaa aagggatgaa gatcctcaca aggtgcca agagatgcct ttccagggtc 23220  
 ctggaaaatc agcatggtgc ccagtggaga tcagcaagac agtcctctgg ccagagagca 23280  
 tcagcgtggt gcgatgtgtg gagttgtttg agggcccggt ggagtgtgag gaggaggagg 23340  
 aggtagagga agaaaaagg agcttctgtg catcgctga gagcagcagg gatgacttcc 23400  
 aggagggaag ggaggcatt gtggcccggc taacagagag cctgttcctg gacctgctg 23460  
 gagaggagaa tgggggcttt tgccagcagg acatggggga gtcattgcctt cttccacctt 23520  
 cgggaagtac gagtgctcac atgccctggg atgagttccc aagtgcaggg cccaaggagg 23580  
 cacctccctg gggcaaggag cagcctctcc acctggagcc aagtcctcct gccagcccga 23640  
 cccagagtcc agacaacctg acttgacag agacgcccct cgtcatcgca ggcaacctg 23700  
 cttaccgcag cttcagcaac tcctgagacc agtcaccgtg tcccagagag ctgggtccag 23760  
 acccactgct ggccagacac ctggagggaag tagaaccga gatgccctgt gtccccagc 23820  
 tctctgagcc aacctgtg ccccaacctg agccagaaac ctgggagcag atcctccgcc 23880  
 gaaatgtcct ccagcatggg gcagctgcag ccccgctctc ggccccacc agtggctatc 23940  
 aggagtttgt acatgcggtg gagcagggtg gcaccagggc cagtgcggtg gtgggcttgg 24000  
 gtccccagag agaggctggt tacaaggcct tctcaagcct gcttgccagc agtgctgtgt 24060

```

ccccagagaa atgtgggttt ggggctagca gtggggaaga ggggtataag cctttccaag 24120
acctcattcc tggctgccct ggggaccctg cccaggtccc tgtccccttg ttcacctttg 24180
gactggacag ggagccacct cgagtcgc agagtcaca tctccaagc agtccccag 24240
agcacctggg tctggagcog ggggaaaagg tagaggacat gccaaagccc cacttcccc 24300
aggagcaggc cacagacccc cttgtggaca gcctgggcag tggcattgtc tactcagccc 24360
ttacctgcca cctgtgcggc cacctgaaac agtgtcatgg ccaggaggat ggtggccaga 24420
cccctgtcat ggceagtcct tgcgtgggt gctgctgtgg agacaggtec tcgcccccta 24480
caacccccct gagggcccca gaccctctc cagggtgggt tccactggag gccagtctgt 24540
gtccggcctc cctggcacc tcgggcattc cagagaagag taaatcctca tcatccttcc 24600
atcctgcccc tggcaatgct cagagctcaa gccagacccc caaaatcgtg aactttgtct 24660
ccgtgggacc cacatacatg aggtctctct aggtgcatgt cctcttgttg ctgagtctgc 24720
agatgaggac tagggcttat ccattgcctgg gaaatgccac ctcttggag gcagccaggc 24780
tggcagattt ccaaaagact tgaagaacca tggatgaag gtgattggcc cactgacgt 24840
tggcctaaca ctgggtgca gagactggac cccgcccagc attgggtgg gctcgccaca 24900
tccatgaga gtagaggga ctgggtgcg gtgcccacg gcaggccct gcaggaaac 24960
tgaggccctt gggcacctcg acttgtgaac gagttgttg ctgctccctc cacagcttct 25020
gcagcagact gtccctgttg taactgcccc aggcattgtt tgcccaccag atcatggccc 25080
acatggaggc ccacctgcct ctgtctcact gaactagaag ccgagcctag aaactaacac 25140
agccatcaag ggaatgactt gggcgccctt ggggaaatcg tgagaattg aacttcaggg 25200
agggtggtca ttgcctagag gtgctcattc atttaacaga gcttccttag gttgatgctg 25260
gaggcagaat cccggtgtc aagggtgtt cagttaaggg gagcaacaga ggacatgaaa 25320
aattgctgtg actaaagcag ggacaatttg ctgccaacaa cccatgcccc gctgtatggc 25380
tgggggtccc tcgtatgcat ggaaccccc gaataaatat gctcagccac cctgtgggcc 25440
gggcaatcca gacagcaggc ataaggcacc agttaccctg catgttggcc cagacctcag 25500
gtgctagggg aggcgggaac cttgggttga gtaatgctcg tctgtgtgtt ttagtttcat 25560
cacctgttat ctgtgtttgc tgaggagagt ggaacagaag ggggtgagtt ttgtataaat 25620
aaagtctctt tgtctcttta ttttttatgt attaaccaaa catacctcca gacactgctg 25680
tgagtgtgtg gtctctgtta actcctggaa ttcaccatc cagaggaacc aggatgcaag 25740
aggttaagaa acttgccgtc tgggtttggg ttcccatac aaggattcaa atagttgatt 25800
taggaagtaa tccgggaaa ccctgctaag gtagtgggga actgaggcag ggaaggacac 25860
aaaccaagaa agtgttacct gaaaggggtc cagatgcaga cccaaaaga gggttcttga 25920
atctcatgca agaaagaatt cagagcgagt ccatagagtc agtgaaagca agttaatgag 25980
gaaagtaaag gaataaaaga atggctactc cgtagacaga gcagccctga gggttgctgg 26040
ctgcctattt ttatggttat tgattaatta tattccaaac aagggttggg ttattatgcc 26100
tcccttttag accatatagg gtaacttct gatgttgcca tggcatttgt aaactgtcat 26160
ggcgtgttg ggagtgtagc agtgaggaca accagaggtc actctgttg ccacttggg 26220
tttgggtggg tagagccatc ttctttactg caacctgttt tatcagcaag gtctttatga 26280
cttgatcgg tgacgacctc ctgtctcatt ctatgactaa gaatgcccta acctcccagg 26340
aatgcagccc agtaagtctc agcctcattt taccagccc ctcttcaaag ctccagttta 26400
aataaacctc tgacaaaagg gtgagttatt caacagatta ccagcatgag taactgatgc 26460
ttacctgccg gggatctctg gaagaccatg catggcacat gccagttat gcctgcaaag 26520
gagagggagc tgggtattt gtccaccagc tccatctgt cattggctga gagctgcttc 26580
caggagcatt aattctccag cacttccagc tactccagga aaaaaaaat tcttcaactg 26640
agagtggag gtgttgagag actctggcac accaagaaga caggacagg acaccaacag 26700
tggctgatga tacactgcca aggtcacaca gctagttagc aacagatcta tagtggaatc 26760
cagacagtgt ctccatcacc caggctctct gtagtgatct gcgcttcaca tccgaggcag 26820
gcagagggat ggtgtgggcc ttagatggga aggtcgggaa cctgaagctc ctatgtctgt 26880
atcacttttg cttctctgag tagctgccct gatttcacac ttgaggggct tggccatttt 26940

```

```

agattccttc ctgctctagg agcctacata ctacactgga aatgatggg agctctctac 27000
ctcacatgca gcctgatgtt tgtagaaac acctccttgc gccaggcatg atggctcatg 27060
gctgtaatcc cagcaatttg ggaggctgag gcgggtgtat cacttgaggt gaggagtcca 27120
agaccagcct ggccaatatg gtgaaaccct atctctacca aaaaataaaa aattagccgg 27180
gtgtggtggt ggggtgcctgt aatctcagct acttgggagg ctgagttggt agaattgctt 27240
caacctggga cgcggagggt gcagtgagct gagattgtgc cattgcactc cagcttggtat 27300
gacagagtga gacctgtct caggaaaaaa aaaaaaaac aaaaaaacc ttgttctaag 27360
ccaaaatcaa tcccttttagc tgcccaaata acacagttaa cagatggaga aacagtttta 27420
gagaggaaaa gggacttgcc caaagtcacc cagagaatgg cagagcctga actagccttc 27480
tggaacttctt gcctccaaaa gctctttata ataaaatata attttaaata aaaatagtta 27540
tctgttttagg gccaaagcaat atgctaagtg ccgtccagcc actgtgtcat ttacgtctcc 27600
aaacagctct agttgggagg ctcaatgatt atcccaattt tacagataag gaaacaggtc 27660
cagagaggtt gaggattagc ctagaaccac acagctagga aatcctggag ccaggatttg 27720
aaccggggtc tgacctaaag gctccagcc gccgtgatat atcagcttat gtcacctga 27780
cacctacgca gatgtcggct cgaatccact ttgcctgagc attgtctcag agaaatctaa 27840
tttaaaaatt aggcagcaaa tagaaaatat atttgactgc tagagatgca atgggactgg 27900
gagcccaaca aaggatctta ggcaaaagaa atccaagttg ttggcctcag caactattac 27960
tgaactggct gggctttggg aagctacaga gggatgagaa gacctggtgg atcagggtggg 28020
cccaactcag gctggccccc accctgcagg aagtaggaaa agtccagggt cataggccca 28080
gtgagatgcc ggctgcggga gtttcagcct ccggggctgg accagagggc aggaggggac 28140
gccctgggt agcagcgcca gagtgggctg agtggcctgg gccctgcgg gggagcttcc 28200
agagatgttg atttgggggt actccctcag ccctgccttt acacagaatt tgtgggggat 28260
gaggggaggg ggaaaggggg gaggaaggca gtgagtgcac ctgaattttt ttttttttt 28320
tacaaaaagt ggcttattgc atttttctga ttactctatc agcacgtgca gaccttttcc 28380
tattcagaga aagcctgaag atataaagag gaaagtgaag aaaaaccacc ggaatccca 28440
tccccgcccc agcatctggc actgtgtggg cgatcacgaa atgagcgctt gtttttgaag 28500
gcgtagtatc tccgtgaaca tccggttgaa caacctttct gactttattt ttcccacgaa 28560
agttattaat taaaaacaa aaagcaaaac accgaaaaaa caaaaaacc agcaagtgtt 28620
tgagctccca ccacgagggg ggctgacgt cactggatcc tcccggcagc cgatgaggct 28680
gcattgggact 28690

```

&lt;210&gt; 2

&lt;211&gt; 2476

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

```

atgggggtggc tttgctctgg gctcctgttc cctgtgagct gcctggctct gctgcagggtg 60
gcaagctctg ggaacatgaa ggtcttgtag gagccacct gcgtctccga ctacatgagc 120
atctctactt gcgagtggaa gatgaatggt ccaccaatt gcagcaccga gctccgcctg 180
ttgtaccagc tggtttttct gctctccgaa gccacacgt gtatccctga gaacaacgga 240
ggcgcggggt gcgtgtgcca cctgctcatg gatgacgtgg tcagtgcgga taactatata 300
ctggacctgt gggctgggca gcagctgctg tggaagggtc cettcaagcc cagcgagcat 360
gtgaaaccca gggccccagg aaacctgaca gttcacacca atgtctccga cactctgctg 420
ctgacctgga gcaaccgcta tccccctgac aattacctgt ataactatct cacctatgca 480
gtcaacattt ggagtgaaaa cgaccgggca gatttcagaa tctataacgt gacctaccta 540
gaaccctccc tccgcatcgc agccagcacc ctgaagtctg ggatttccta cagggcacgg 600

```

```

gtgagggcct gggctcagtg ctataacacc acctggagtg agtggagccc cagcaccaag 660
tggcacaact cctacagggg gcccttcgag cagcacctcc tgctgggCGT cagcgtttcc 720
tgcatgtgca tcttgccgtg ctgctgttg tgctatgtca gcatcaccaa gattaagaaa 780
gaatgggtgg atcagattcc caaccagcc cgcagccgcc tcgtggctat aataatccag 840
gatgctcagg ggtcacagtg ggagaagcgg tcccagggcc aggaaccagc caagtGCCA 900
cactggaaga attgtcttac caagctcttg ccctgttttc tggagcaca catgaaaagg 960
gatgaagatc ctcacaaggc tgccaaagag atgcctttcc agggctctgg aaaatcagca 1020
tggtgcccag tggagatcag caagacagtc ctctggccag agagcatcag cgtggtgcga 1080
tgtgtggagt tgtttgaggc cccggtggag tgtgaggagg aggaggagg agaggaagaa 1140
aaagggagct tctgtgcatc gcctgagagc agcagggatg acttccagga gggaaggagg 1200
ggcattgtgg cccggctaac agagagcctg ttcctggacc tgctcggaga ggagaatggg 1260
ggcttttgcc agcaggacat gggggagtc tgccctcttc caccttcggg aagtacgagt 1320
gtcacatgc cctgggatga gttcccaagt gcaggGCCA aggaggcacc tccctggggc 1380
aaggagcagc ctctccacct ggagccaagt cctcctgcca gcccgaccca gagtccagac 1440
aacctgactt gcacagagac gcccctcgtc atcgcaggca accctgctta ccgcagcttc 1500
agcaactccc tgagccagtc accgtgtccc agagagctgg gtccagacc actgctggcc 1560
agacacctgg aggaagtaga acccgagatg ccctgtgtcc ccagctctc tgagccaacc 1620
actgtgcccc aacctgagcc agaaacctgg gacagatcc tccgccgaaa tgcctccag 1680
catggggcag ctgcagcccc cgtctcgcc cccaccagt gctatcagga gtttgtacat 1740
gcggtggagc aggggtggc caggccagt gcggtggtg gcttgggtcc cccaggagag 1800
gctggttaca aggccttctc aagcctgctt gccagcagt ctgtgtcccc agagaaatgt 1860
gggtttgggg ctgacagtgg ggaagagggg tataagcctt tccaagacct cattcctggc 1920
tgccctgggg accctgcccc agtccctgtc ccctgttca ccttggact ggacagggag 1980
ccacctcgca gtccgagag ctacatctc ccaagcagct cccagagca cctgggtctg 2040
gagccggggg aaaaggtaga ggacatgcca aagccccac tccccagga gcaggccaca 2100
gaccccttg tggacagcct gggcagtggc attgtctact cagcccttac ctgccacctg 2160
tgcgccacc tgaacagtg tcatggccag gaggatggtg gccagacccc tgtcatggcc 2220
agtcttgct gtggtgctg ctgtggagac aggtcctcgc cccctacaac cccctgagg 2280
gccccagacc cctctccagg tggggttcca ctggaggcca gtctgtgtcc ggcctccctg 2340
gcaccctcgg gcatctcaga gaagagtaaa tctcatcat ccttccatcc tgccctggc 2400
aatgtcaga gctcaagcca gacccccaaa atcgtgaact ttgtctccgt gggaccacaa 2460
tacatgaggg tctctt 2476

```

&lt;210&gt; 3

&lt;211&gt; 825

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

```

Met Gly Trp Leu Cys Ser Gly Leu Leu Phe Pro Val Ser Cys Leu Val
  1             5             10            15

```

```

Leu Leu Gln Val Ala Ser Ser Gly Asn Met Lys Val Leu Gln Glu Pro
      20             25             30

```

```

Thr Cys Val Ser Asp Tyr Met Ser Ile Ser Thr Cys Glu Trp Lys Met
    35             40             45

```

Asn Gly Pro Thr Asn Cys Ser Thr Glu Leu Arg Leu Leu Tyr Gln Leu  
 50 55 60

Val Phe Leu Leu Ser Glu Ala His Thr Cys Ile Pro Glu Asn Asn Gly  
 65 70 75 80

Gly Ala Gly Cys Val Cys His Leu Leu Met Asp Asp Val Val Ser Ala  
 85 90 95

Asp Asn Tyr Thr Leu Asp Leu Trp Ala Gly Gln Gln Leu Leu Trp Lys  
 100 105 110

Gly Ser Phe Lys Pro Ser Glu His Val Lys Pro Arg Ala Pro Gly Asn  
 115 120 125

Leu Thr Val His Thr Asn Val Ser Asp Thr Leu Leu Leu Thr Trp Ser  
 130 135 140

Asn Pro Tyr Pro Pro Asp Asn Tyr Leu Tyr Asn His Leu Thr Tyr Ala  
 145 150 155 160

Val Asn Ile Trp Ser Glu Asn Asp Pro Ala Asp Phe Arg Ile Tyr Asn  
 165 170 175

Val Thr Tyr Leu Glu Pro Ser Leu Arg Ile Ala Ala Ser Thr Leu Lys  
 180 185 190

Ser Gly Ile Ser Tyr Arg Ala Arg Val Arg Ala Trp Ala Gln Cys Tyr  
 195 200 205

Asn Thr Thr Trp Ser Glu Trp Ser Pro Ser Thr Lys Trp His Asn Ser  
 210 215 220

Tyr Arg Glu Pro Phe Glu Gln His Leu Leu Leu Gly Val Ser Val Ser  
 225 230 235 240

Cys Ile Val Ile Leu Ala Val Cys Leu Leu Cys Tyr Val Ser Ile Thr  
 245 250 255

Lys Ile Lys Lys Glu Trp Trp Asp Gln Ile Pro Asn Pro Ala Arg Ser  
 260 265 270

Arg Leu Val Ala Ile Ile Ile Gln Asp Ala Gln Gly Ser Gln Trp Glu  
 275 280 285

Lys Arg Ser Arg Gly Gln Glu Pro Ala Lys Cys Pro His Trp Lys Asn  
 290 295 300

Cys Leu Thr Lys Leu Leu Pro Cys Phe Leu Glu His Asn Met Lys Arg  
 305 310 315 320  
 Asp Glu Asp Pro His Lys Ala Ala Lys Glu Met Pro Phe Gln Gly Ser  
 325 330 335  
 Gly Lys Ser Ala Trp Cys Pro Val Glu Ile Ser Lys Thr Val Leu Trp  
 340 345 350  
 Pro Glu Ser Ile Ser Val Val Arg Cys Val Glu Leu Phe Glu Ala Pro  
 355 360 365  
 Val Glu Cys Glu Glu Glu Glu Glu Val Glu Glu Glu Lys Gly Ser Phe  
 370 375 380  
 Cys Ala Ser Pro Glu Ser Ser Arg Asp Asp Phe Gln Glu Gly Arg Glu  
 385 390 395 400  
 Gly Ile Val Ala Arg Leu Thr Glu Ser Leu Phe Leu Asp Leu Leu Gly  
 405 410 415  
 Glu Glu Asn Gly Gly Phe Cys Gln Gln Asp Met Gly Glu Ser Cys Leu  
 420 425 430  
 Leu Pro Pro Ser Gly Ser Thr Ser Ala His Met Pro Trp Asp Glu Phe  
 435 440 445  
 Pro Ser Ala Gly Pro Lys Glu Ala Pro Pro Trp Gly Lys Glu Gln Pro  
 450 455 460  
 Leu His Leu Glu Pro Ser Pro Pro Ala Ser Pro Thr Gln Ser Pro Asp  
 465 470 475 480  
 Asn Leu Thr Cys Thr Glu Thr Pro Leu Val Ile Ala Gly Asn Pro Ala  
 485 490 495  
 Tyr Arg Ser Phe Ser Asn Ser Leu Ser Gln Ser Pro Cys Pro Arg Glu  
 500 505 510  
 Leu Gly Pro Asp Pro Leu Leu Ala Arg His Leu Glu Glu Val Glu Pro  
 515 520 525  
 Glu Met Pro Cys Val Pro Gln Leu Ser Glu Pro Thr Thr Val Pro Gln  
 530 535 540  
 Pro Glu Pro Glu Thr Trp Glu Gln Ile Leu Arg Arg Asn Val Leu Gln  
 545 550 555 560

His Gly Ala Ala Ala Ala Pro Val Ser Ala Pro Thr Ser Gly Tyr Gln  
 565 570 575

Glu Phe Val His Ala Val Glu Gln Gly Gly Thr Gln Ala Ser Ala Val  
 580 585 590

Val Gly Leu Gly Pro Pro Gly Glu Ala Gly Tyr Lys Ala Phe Ser Ser  
 595 600 605

Leu Leu Ala Ser Ser Ala Val Ser Pro Glu Lys Cys Gly Phe Gly Ala  
 610 615 620

Ser Ser Gly Glu Glu Gly Tyr Lys Pro Phe Gln Asp Leu Ile Pro Gly  
 625 630 635 640

Cys Pro Gly Asp Pro Ala Pro Val Pro Val Pro Leu Phe Thr Phe Gly  
 645 650 655

Leu Asp Arg Glu Pro Pro Arg Ser Pro Gln Ser Ser His Leu Pro Ser  
 660 665 670

Ser Ser Pro Glu His Leu Gly Leu Glu Pro Gly Glu Lys Val Glu Asp  
 675 680 685

Met Pro Lys Pro Pro Leu Pro Gln Glu Gln Ala Thr Asp Pro Leu Val  
 690 695 700

Asp Ser Leu Gly Ser Gly Ile Val Tyr Ser Ala Leu Thr Cys His Leu  
 705 710 715 720

Cys Gly His Leu Lys Gln Cys His Gly Gln Glu Asp Gly Gly Gln Thr  
 725 730 735

Pro Val Met Ala Ser Pro Cys Cys Gly Cys Cys Cys Gly Asp Arg Ser  
 740 745 750

Ser Pro Pro Thr Thr Pro Leu Arg Ala Pro Asp Pro Ser Pro Gly Gly  
 755 760 765

Val Pro Leu Glu Ala Ser Leu Cys Pro Ala Ser Leu Ala Pro Ser Gly  
 770 775 780

Ile Ser Glu Lys Ser Lys Ser Ser Ser Ser Phe His Pro Ala Pro Gly  
 785 790 795 800

Asn Ala Gln Ser Ser Ser Gln Thr Pro Lys Ile Val Asn Phe Val Ser  
 805 810 815



Val Gly Pro Thr Tyr Met Arg Val Ser  
820 825

<210> 4  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 4  
ttgcaccact gcact 15

<210> 5  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 5  
ttgcaccgct gcact 15

<210> 6  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 6  
ttttgtgcta ttccc 15

<210> 7  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 7  
ttttgtgtta ttccc 15

<210> 8  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 8  
ctgggccgct caggc 15

<210> 9  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 9  
ctgggcctct caggc

15

<210> 10  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 10  
taagcctgcg ctgga

15

<210> 11  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 11  
taagcctccg ctgga

15

<210> 12  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 12  
agaacaacgg aggcg

15

<210> 13  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 13  
agaacaatgg aggcg

15

<210> 14

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 14  
cggaggcgcg ggggtg.

15

<210> 15  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 15  
cggaggcacg ggggtg

15

<210> 16  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 16  
gtgcggataa ctata

15

<210> 17  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 17  
gtgcggacaa ctata

15

<210> 18  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 18  
cggagtgcgg caggg

15

<210> 19  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 19

cggagtgtgg caggg

15

&lt;210&gt; 20

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 20

gcctgggctg aggg

15

&lt;210&gt; 21

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

gcctgggttg aggg

15

&lt;210&gt; 22

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

tgggtgggc aggg

15

&lt;210&gt; 23

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 23

tgggtgagc aggg

15

&lt;210&gt; 24

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 24

ttctccgca gtgaa

15

<210> 25  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 25  
ttctcccaca gtgaa

15

<210> 26  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 26  
gtgaaaacga cccgg

15

<210> 27  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 27  
gtgaaatga cccgg

15

<210> 28  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 28  
ggcaagccct ggggc

15

<210> 29  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 29  
ggcaagctct ggggc

15

<210> 30

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 30  
gccctggggc tggat

15

<210> 31  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 31  
gccctggagc tggat

15

<210> 32  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 32  
atagcaaadc ccagg

15

<210> 33  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 33  
atagcaattc ccagg

15

<210> 34  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 34  
gctctgccct aggca

15

<210> 35  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 35

gctctgcact aggca

15

&lt;210&gt; 36

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 36

ccccacccc tcaca

15

&lt;210&gt; 37

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 37

ccccactcc tcaca

15

&lt;210&gt; 38

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 38

tcctccgca tcgca

15

&lt;210&gt; 39

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

tcctccaca tcgca

15

&lt;210&gt; 40

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

cacctgctgt ggtgt

15

<210> 41  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 41  
cacctgccgt ggtgt

15

<210> 42  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 42  
atgtctgaag tagac

15

<210> 43  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 43  
atgtctgcag tagac

15

<210> 44  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 44  
tgaccaacct ttgct

15

<210> 45  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 45  
tgaccaactt ttgct

15

<210> 46



<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 46  
cctgttttct ggagc

15

<210> 47  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 47  
cctgtttcct ggagc

15

<210> 48  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 48  
tggacctgct cggag

15

<210> 49  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 49  
tggaccttct cggag

15

<210> 50  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 50  
agtcatgcct tcttc

15

<210> 51  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 51  
agtcattgtct tcttc

15

<210> 52  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 52  
gccttcttcc acctt

15

<210> 53  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 53  
gccttctccc acctt

15

<210> 54  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 54  
cagccccgt ctcgg

15

<210> 55  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 55  
cagcccctgt ctcgg

15

<210> 56  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 56  
ggagtttgta catgc

15

<210> 57  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 57  
ggagtttata catgc

15

<210> 58  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 58  
cagctcccca gagca

15

<210> 59  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 59  
cagctcctca gagca

15

<210> 60  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 60  
agacaggtcc tcgcc

15

<210> 61  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 61  
agacagggcc tcgcc

15

<210> 62

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 62  
ctgcccctgg caatg

15

<210> 63  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 63  
ctgcccccg caatg

15

<210> 64  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 64  
aggtgcatgt cctct

15

<210> 65  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 65  
aggtgcacgt cctct

15

<210> 66  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 66  
gtgcatgtcc tcttg

15

<210> 67  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 67

gtgcatgccc tcttg

15

&lt;210&gt; 68

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 68

ggcttatcca tgcct

15

&lt;210&gt; 69

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 69

ggcttattca tgcct

15

&lt;210&gt; 70

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 70

agccaggctg gcaga

15

&lt;210&gt; 71

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 71

agccagggtg gcaga

15

&lt;210&gt; 72

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 72

ggcccacatg gaggc

15

<210> 73  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 73  
ggcccacgtg gaggc

15

<210> 74  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 74  
taacacagcc atcaa

15

<210> 75  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 75  
taacacaacc atcaa

15

<210> 76  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 76  
taatgctcgt ctgtg

15

<210> 77  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 77  
taatgcttgt ctgtg

15

<210> 78

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 78  
acttgccgtc tgggt

15

<210> 79  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 79  
acttgccatc tgggt

15

<210> 80  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 80  
ctgagattgc accac

15

<210> 81  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 81  
ggctggagtg cagtg

15

<210> 82  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 82  
ctgagattgc accgc

15

<210> 83  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 83

ggctggagtg cagcg

15

&lt;210&gt; 84

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 84

ctgtgctttt gtgct

15

&lt;210&gt; 85

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 85

accaagggga atagc

15

&lt;210&gt; 86

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 86

ctgtgctttt gtgtt

15

&lt;210&gt; 87

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 87

accaagggga ataac

15

&lt;210&gt; 88

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 88

gagttcctgg gccgc

15



<210> 89  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 89  
ggagcagcct gagcg

15

<210> 90  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 90  
gagttcctgg gcctc

15

<210> 91  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 91  
ggagcagcct gagag

15

<210> 92  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 92  
tccgagtaag cctgc

15

<210> 93  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 93  
tccagctcca gcgca

15

<210> 94

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 94  
tccgagtaag cctcc

15

<210> 95  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 95  
tccagctcca gcgga

15

<210> 96  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 96  
tccctgagaa caacg

15

<210> 97  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 97  
accccgcgcc tccgt

15

<210> 98  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 98  
tccctgagaa caatg

15

<210> 99  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 99  
accccgcgcc tccat

15

<210> 100  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 100  
gaacaacgga ggcgc

15

<210> 101  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 101  
cacacgcacc ccgcg

15

<210> 102  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 102  
gaacaacgga ggcac

15

<210> 103  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 103  
cacacgcacc ccgtg

15

<210> 104  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 104  
tggtcagtgc ggata

15

<210> 105

<211> 15

<212> DNA

<213> Homo sapiens

<400> 105

ccagtgtata gttat

15

<210> 106

<211> 15

<212> DNA

<213> Homo sapiens

<400> 106

gcagggcgga gtgcg

15

<210> 107

<211> 15

<212> DNA

<213> Homo sapiens

<400> 107

agccaccct gccgc

15

<210> 108

<211> 15

<212> DNA

<213> Homo sapiens

<400> 108

gcagggcgga gtgtg

15

<210> 109

<211> 15

<212> DNA

<213> Homo sapiens

<400> 109

agccaccct gccac

15

<210> 110

<211> 15

<212> DNA

<213> Homo sapiens

<400> 110

acagctgcct gggct

15

<210> 111

<211> 15

<212> DNA

<213> Homo sapiens

<400> 111

caccccaccc tcagc

15

<210> 112

<211> 15

<212> DNA

<213> Homo sapiens

<400> 112

acagctgcct gggtt

15

<210> 113

<211> 15

<212> DNA

<213> Homo sapiens

<400> 113

acagctgcct gggtt

15

<210> 114

<211> 15

<212> DNA

<213> Homo sapiens

<400> 114

tgaggggtggg gtggg

15

<210> 115

<211> 15

<212> DNA

<213> Homo sapiens

<400> 115  
cctcctcccc tgccc

15

<210> 116  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 116  
tgaggggtggg gtgag

15

<210> 117  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 117  
cctcctcccc tgctc

15

<210> 118  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 118  
ggccgcttct cccgc

15

<210> 119  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 119  
ctgggtttca ctgcg

15

<210> 120  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 120  
ggccgcttct ccac

15

<210> 121  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 121  
ctgggtttca ctgtg

15

<210> 122  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 122  
tttgagtgaaacg

15

<210> 123  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 123  
catctgccgg gtcgt

15

<210> 124  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 124  
tttgagtgaaatg

15

<210> 125  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 125  
catctgccgg gtcac

15

<210> 126

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 126  
ctgggaggca agccc

15

<210> 127  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 127  
tatccagccc caggg

15

<210> 128  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 128  
tggtcagtgc ggaca

15

<210> 129  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 129  
ccagtggtata gttgt

15

<210> 130  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 130  
ctgggaggca agctc

15

<210> 131  
<211> 15  
<212> DNA  
<213> Homo sapiens



<400> 131  
tatccagccc cagag

15

<210> 132  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 132  
aggcaagccc tgggg

15

<210> 133  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 133  
tttgctatcc agccc

15

<210> 134  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 134  
aggcaagccc tggag

15

<210> 135  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 135  
tttgctatcc agctc

15

<210> 136  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 136  
ggctggatag caaat

15

<210> 137  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 137  
ctagctcctg ggatt

15

<210> 138  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 138  
ggctggatag caatt

15

<210> 139  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 139  
ctagctcctg ggaat

15

<210> 140  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 140  
cacctggctc tgccc

15

<210> 141  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 141  
gggacttgcc taggg

15

<210> 142

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 142  
cacctggctc tgcac

15

<210> 143  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 143  
gggacttgcc tagtg

15

<210> 144  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 144  
cctggcccc ccccc

15

<210> 145  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 145  
ctctgatgtg agggg

15

<210> 146  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 146  
cctggcccc cactc

15

<210> 147  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 147  
ctctgatgtg aggag

15

<210> 148  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 148  
gaaccctccc tccgc

15

<210> 149  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 149  
gctggctgcg atgcg

15

<210> 150  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 150  
gaaccctccc tccac

15

<210> 151  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 151  
gctggctgcg atgtg

15

<210> 152  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 152  
tagatacacc tgctg

15

<210> 153  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 153  
gcagatacac cacag 15

<210> 154  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 154  
tagatacacc tgccg 15

<210> 155  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 155  
gcagatacac cacgg 15

<210> 156  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 156  
gaaggcatgt ctgaa 15

<210> 157  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 157  
atggctgtct acttc 15

<210> 158

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 158  
gaaggcatgt ctgca

15

<210> 159  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 159  
atggctgtct actgc

15

<210> 160  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 160  
gaaccctgac caacc

15

<210> 161  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 161  
tgcaaaagca aagga

15

<210> 162  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 162  
gaaccctgac caatc

15

<210> 163  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 163  
tgcaaaagca aagaa

15

<210> 164  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 164  
tcttgccctg ttttc

15

<210> 165  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 165  
tgttgtgctc cagaa

15

<210> 166  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 166  
tcttgccctg tttcc

15

<210> 167  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 167  
tgttgtgctc cagga

15

<210> 168  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 168  
tgttcctgga cctgc

15

<210> 169  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 169  
tctcctctcc gagca

15

<210> 170  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 170  
tggtcctgga ccttc

15

<210> 171  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 171  
tctcctctcc gagaa

15

<210> 172  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 172  
tgggggagtc atgcc

15

<210> 173  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 173  
aagtggaag aaggc

15

<210> 174



<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 174  
tgggggagtc atgtc

15

<210> 175  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 175  
aaggtggaag aagac

15

<210> 176  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 176  
agtcatgcct tcttc

15

<210> 177  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 177  
ttcccgaagg tggaa

15

<210> 178  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 178  
agtcatgcct tctcc

15

<210> 179  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 179

ttcccgaagg tggga

15

&lt;210&gt; 180

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 180

cagctgcagc ccccg

15

&lt;210&gt; 181

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 181

tgggggccga gacgg

15

&lt;210&gt; 182

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 182

cagctgcagc ccctg

15

&lt;210&gt; 183

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 183

tgggggccga gacag

15

&lt;210&gt; 184

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 184

ctatcaggag tttgt

15

<210> 185  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 185  
tccaccgcat gtaca

15

<210> 186  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 186  
ctatcaggag tttat

15

<210> 187  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 187  
tccaaccgcat gtata

15

<210> 188  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 188  
cccaagcagc tcccc

15

<210> 189  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 189  
cccagtgct ctggg

15

<210> 190

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 190  
cccaagcagc tcctc

15

<210> 191  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 191  
cccaggtgct ctgag

15

<210> 192  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 192  
ctgtggagac aggtc

15

<210> 193  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 193  
gtagggggcg aggac

15

<210> 194  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 194  
ctgtggagac agggc

15

<210> 195  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 195  
gtagggggcg aggcc

15

<210> 196  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 196  
tccatcctgc ccctg

15

<210> 197  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 197  
tctgagcatt gccag

15

<210> 198  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 198  
tccatcctgc ccccg

15

<210> 199  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 199  
tctgagcatt gccgg

15

<210> 200  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 200  
tctcttaggt gcatg

15

<210> 201  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 201  
gcaacaagag gacat

15

<210> 202  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 202  
tctcttaggt gcacg

15

<210> 203  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 203  
gcaacaagag gacgt

15

<210> 204  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 204  
tcttaggtgc atgtc

15

<210> 205  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 205  
cagcaacaag aggac

15

<210> 206

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 206  
tcttaggtgc atgcc

15

<210> 207  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 207  
cagcaacaag agggc

15

<210> 208  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 208  
gactagggct tatcc

15

<210> 209  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 209  
tttcccaggc atgga

15

<210> 210  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 210  
gactagggct tattc

15

<210> 211  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 211  
tttcccaggc atgaa

15

<210> 212  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 212  
gaaggcagcc aggct

15

<210> 213  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 213  
tggaaatctg ccagc

15

<210> 214  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 214  
gaaggcagcc agggt

15

<210> 215  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 215  
tggaaatctg ccacc

15

<210> 216  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 216  
gatcatggcc cacat

15



<210> 217  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 217  
aggtgggcct ccatg 15

<210> 218  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 218  
gatcatggcc cacgt 15

<210> 219  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 219  
aggtgggcct ccacg 15

<210> 220  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 220  
agaaactaac acagc 15

<210> 221  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 221  
attcccttga tggct 15

<210> 222

<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 222  
agaaactaac acaac

15

<210> 223  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 223  
attcccttga tgggt

15

<210> 224  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 224  
gttgagtaat gctcg

15

<210> 225  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 225  
aaaacacaca gacga

15

<210> 226  
<211> 15  
<212> DNA  
<213> Homo sapiens

<400> 226  
gttgagtaat gcttg

15

<210> 227  
<211> 15  
<212> DNA  
<213> Homo sapiens

&lt;400&gt; 227

aaaacacaca gacaa

15

&lt;210&gt; 228

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 228

taagaaactt gccgt

15

&lt;210&gt; 229

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 229

acccaaacc agacg

15

&lt;210&gt; 230

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 230

taagaaactt gccat

15

&lt;210&gt; 231

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 231

acccaaacc agatg

15

&lt;210&gt; 232

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 232

agattgcacc

10

<210> 233  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 233  
tggagtgcag

10

<210> 234  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 234  
tgcttttgtg

10

<210> 235  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 235  
aaggggaata

10

<210> 236  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 236  
ttcctgggcc

10

<210> 237  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 237  
gcagcctgag

10

<210> 238

<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 238  
gagtaagcct

10

<210> 239  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 239  
agctccagcg

10

<210> 240  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 240  
ctgagaacaa

10

<210> 241  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 241  
ccgcgcctcc

10

<210> 242  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 242  
caacggaggc

10

<210> 243  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 243  
acgcacccccg

10

<210> 244  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 244  
tcagtgcgga

10

<210> 245  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 245  
gtgtatagtt

10

<210> 246  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 246  
gggcggagtg

10

<210> 247  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 247  
caccctgcc

10

<210> 248  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 248  
gctgcctggg

10

<210> 249

<211> 10

<212> DNA

<213> Homo sapiens

<400> 249

cccaccctca

10

<210> 250

<211> 10

<212> DNA

<213> Homo sapiens

<400> 250

gggtggggtg

10

<210> 251

<211> 10

<212> DNA

<213> Homo sapiens

<400> 251

cctcccctgc

10

<210> 252

<211> 10

<212> DNA

<213> Homo sapiens

<400> 252

cgcttctccc

10

<210> 253

<211> 10

<212> DNA

<213> Homo sapiens

<400> 253

ggtttcactg

10

<210> 254

<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 254  
ggagtgaaaa

10

<210> 255  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 255  
ctgccggggtc

10

<210> 256  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 256  
ggaggcaagc

10

<210> 257  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 257  
ccagccccag

10

<210> 258  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 258  
caagccctgg

10

<210> 259  
<211> 10  
<212> DNA  
<213> Homo sapiens



<400> 259  
gctatccagc

10

<210> 260  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 260  
tggatagcaa

10

<210> 261  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 261  
gctcctggga

10

<210> 262  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 262  
ctggctctgc

10

<210> 263  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 263  
acttgccctag

10

<210> 264  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 264  
ggccccccac

10

<210> 265  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 265  
tgatgtgagg

10

<210> 266  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 266  
ccctccctcc

10

<210> 267  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 267  
ggctgcatg

10

<210> 268  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 268  
atacacctgc

10

<210> 269  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 269  
gatacaccac

10

<210> 270

<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 270  
ggcatgtctg

10

<210> 271  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 271  
gctgtctact

10

<210> 272  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 272  
ccctgaccaa

10

<210> 273  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 273  
aaaagcaaag

10

<210> 274  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 274  
tgccctgttt

10

<210> 275  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 275  
tgtgctccag

10

<210> 276  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 276  
tcttgacct

10

<210> 277  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 277  
cctctccgag

10

<210> 278  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 278  
gggagtcag

10

<210> 279  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 279  
gtggaagaag

10

<210> 280  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 280  
catgccttct

10

<210> 281  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 281  
ccgaaggtgg

10

<210> 282  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 282  
ctgcagcccc

10

<210> 283  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 283  
gggccgagac

10

<210> 284  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 284  
tcaggagttt

10

<210> 285  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 285  
accgcatgta

10

<210> 286

<211> 10

<212> DNA

<213> Homo sapiens

<400> 286

aagcagctcc

10

<210> 287

<211> 10

<212> DNA

<213> Homo sapiens

<400> 287

aggtgctctg

10

<210> 288

<211> 10

<212> DNA

<213> Homo sapiens

<400> 288

tggagacagg

10

<210> 289

<211> 10

<212> DNA

<213> Homo sapiens

<400> 289

gggggcgagg

10

<210> 290

<211> 10

<212> DNA

<213> Homo sapiens

<400> 290

atcctgcccc

10

<210> 291

<211> 10

<212> DNA

<213> Homo sapiens

<400> 291  
gagcattgcc

10

<210> 292  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 292  
cttaggtgca

10

<210> 293  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 293  
acaagaggac

10

<210> 294  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 294  
taggtgcatg

10

<210> 295  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 295  
caacaagagg

10

<210> 296  
<211> 10  
<212> DNA  
<213> Homo sapiens

<400> 296  
tagggcttat

10

<210> 297

<211> 10

<212> DNA

<213> Homo sapiens

<400> 297

cccaggcatg

10

<210> 298

<211> 10

<212> DNA

<213> Homo sapiens

<400> 298

ggcagccagg

10

<210> 299

<211> 10

<212> DNA

<213> Homo sapiens

<400> 299.

aaatctgcca

10

<210> 300

<211> 10

<212> DNA

<213> Homo sapiens

<400> 300

catggccac

10

<210> 301

<211> 10

<212> DNA

<213> Homo sapiens

<400> 301

tgggcctcca

10

<210> 302



&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 302

aactaacaca

10

&lt;210&gt; 303

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 303

cccttgatgg

10

&lt;210&gt; 304

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 304

gagtaatgct

10

&lt;210&gt; 305

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 305

acacacagac

10

&lt;210&gt; 306

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 306

gaaacttgcc

10

&lt;210&gt; 307

&lt;211&gt; 10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 307  
caaaccaga

10

<210> 308  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 308  
ccacagtcac cccgacacta gc

22

<210> 309  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 309  
tattccagcc gtatccatgt gc

22

<210> 310  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 310  
ccttggtgca tgtggtaaga gg

22

<210> 311  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 311  
tttcaaaggt gggaggactg agg

23

<210> 312  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 312  
gcagtgaact gggattgtgc

20

<210> 313

<211> 23

<212> DNA

<213> Homo sapiens

<400> 313

aactcccctt ctctgatgtg agg

23

<210> 314

<211> 23

<212> DNA

<213> Homo sapiens

<400> 314

tcacagttac agaggtggca agc

23

<210> 315

<211> 22

<212> DNA

<213> Homo sapiens

<400> 315

ctgcctacct ggcagataca cc

22

<210> 316

<211> 22

<212> DNA

<213> Homo sapiens

<400> 316

agctgtcact ccacctcctt gg

22

<210> 317

<211> 24

<212> DNA

<213> Homo sapiens

<400> 317

aaagcctctg gtctgctaata gacc

24

<210> 318

<211> 22

<212> DNA

<213> Homo sapiens

<400> 318

gggaggagat tcagagcact cc

22

<210> 319

<211> 22

<212> DNA

<213> Homo sapiens

<400> 319

cagtccacgt ttccagaaca cc

22

<210> 320

<211> 21

<212> DNA

<213> Homo sapiens

<400> 320

ggcttgggat aatggtgttg c

21

<210> 321

<211> 23

<212> DNA

<213> Homo sapiens

<400> 321

tacttcccga aggtggaaga agg

23

<210> 322

<211> 24

<212> DNA

<213> Homo sapiens

<400> 322

cagtggagat cagcaagaca gtcc

24

<210> 323

<211> 22

<212> DNA

<213> Homo sapiens

&lt;400&gt; 323

gggcacatctcg ggttctactt cc

22

&lt;210&gt; 324

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 324

gggaagtacg agtgctcaca tgc

23

&lt;210&gt; 325

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 325

cttataacccc tcttccccac tgc

23

&lt;210&gt; 326

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 326

tctctgagcc aaccactgtg c

21

&lt;210&gt; 327

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 327

ggctgagtag acaatgccac tgc

23

&lt;210&gt; 328

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 328

ctgtgtcccc agagaaatgt gg

22

<210> 329

<211> 23

<212> DNA

<213> Homo sapiens

<400> 329

gactcagcaa caagaggaca tgc

23

<210> 330

<211> 24

<212> DNA

<213> Homo sapiens

<400> 330

ggcattgtct actcagccct tacc

24

<210> 331

<211> 20

<212> DNA

<213> Homo sapiens

<400> 331

acaagtcgag gtgcccaagg

20

<210> 332

<211> 25

<212> DNA

<213> Homo sapiens

<400> 332

cccacatata tgagggtctc ttagg

25

<210> 333

<211> 21

<212> DNA

<213> Homo sapiens

<400> 333

attctgcctc cagcatcaac c

21

<210> 334

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 334

aacagagctt ccttaggttg atgc

24

&lt;210&gt; 335

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 335

cctcagttcc ccactacctt agc

23

&lt;210&gt; 336

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 336

gcgctggccc tcaacttt

18

&lt;210&gt; 337

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 337

gtccctggag atgggacctc

20

&lt;210&gt; 338

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 338

gccccagat ctgtcctcac

20

&lt;210&gt; 339

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 339

ggaaaataca ggcggcttcc

20

&lt;210&gt; 340

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 340

ggctctgaat ctgtgtggtg ct

22

&lt;210&gt; 341

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 341

agccaggtga gaagccaggt

20

&lt;210&gt; 342

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 342

ggcctgaaca ggacgaaca

20

&lt;210&gt; 343

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 343

ggcaggattg ccattagagg

20

&lt;210&gt; 344

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 344

tgagtcagtg gtttgacctc ca

22



<210> 345

<211> 20

<212> DNA

<213> Homo sapiens

<400> 345

gcctctgtct cccctgcaac

20

<210> 346

<211> 20

<212> DNA

<213> Homo sapiens

<400> 346

ccacttttgc catcgaccac

20

<210> 347

<211> 20

<212> DNA

<213> Homo sapiens

<400> 347

ctgccgtccc ttgaaggcta

20

<210> 348

<211> 22

<212> DNA

<213> Homo sapiens

<400> 348

ccctaccctc agggatttct ca

22

<210> 349

<211> 20

<212> DNA

<213> Homo sapiens

<400> 349

cccattctcc tctccgagca

20

<210> 350

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 350

atcagcgtgg tgcgatgtgt

20

&lt;210&gt; 351

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 351

acccagctct ctgggacacg

20

&lt;210&gt; 352

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 352

gggatgagtt cccaagtga g

21

&lt;210&gt; 353

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 353

tggcaagcag gcttgagaag

20

&lt;210&gt; 354

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 354

ccccaacctg agccagaaac

20

&lt;210&gt; 355

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 355

tgtccacaag ggggtctgtg

20

&lt;210&gt; 356

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 356

ggctagcagt ggggaagagg

20

&lt;210&gt; 357

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 357

attgccaggg gcaggatg

18

&lt;210&gt; 358

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 358

ccctgtcatg gccagtcctt

20

&lt;210&gt; 359

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 359

gcgacccagt gccctctact

20

&lt;210&gt; 360

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 360

tgcattgcct cttgttgctg ag

22

<210> 361

<211> 20

<212> DNA

<213> Homo sapiens

<400> 361

caatgaccac cctccctgaa

20

<210> 362

<211> 19

<212> DNA

<213> Homo sapiens

<400> 362

cggtgtcaa ggggtgttc

19

<210> 363

<211> 20

<212> DNA

<213> Homo sapiens

<400> 363

ccaaaccag acggcaagtt

20

<210> 364

<211> 24

<212> DNA

<213> Homo sapiens

<400> 364

aaacccttg actccaagtg atcc

24

<210> 365

<211> 24

<212> DNA

<213> Homo sapiens

<400> 365

aagcgattct tctgcctcag cctc

24

<210> 366

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 366

ggacagttgt tgtgtagctc accc

24

&lt;210&gt; 367

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 367

ctatgttgcc caagctgacc tc

22

&lt;210&gt; 368

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 368

gccctcaact ttgcctgcac

20

&lt;210&gt; 369

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 369

agtccagtat tccagccgta tcc

23

&lt;210&gt; 370

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 370

tgatcgggaa gctggaagag tc

22

&lt;210&gt; 371

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 371  
cgtttcaaag gtgggaggac tg

22

<210> 372  
<211> 21  
<212> DNA  
<213> Homo sapiens

<400> 372  
cgaccaaaaa tctgggtggt g

21

<210> 373  
<211> 21  
<212> DNA  
<213> Homo sapiens

<400> 373  
caggaagcaa agggacttgc c

21

<210> 374  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 374  
tcttaaacad ggtggggtca gc

22

<210> 375  
<211> 21  
<212> DNA  
<213> Homo sapiens

<400> 375  
catggaaatt gtgggcttgt g

21

<210> 376  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 376  
atgtgcaaga gggagagtgg tg

22

<210> 377

<211> 22

<212> DNA

<213> Homo sapiens

<400> 377

tgactgagag gactgcaaag gg

22

<210> 378

<211> 24

<212> DNA

<213> Homo sapiens

<400> 378

gcctgatctc tgatgccaaa taag

24

<210> 379

<211> 20

<212> DNA

<213> Homo sapiens

<400> 379

tttgccattc cagaagccag

20

<210> 380

<211> 23

<212> DNA

<213> Homo sapiens

<400> 380

gatctgtgtg atgtcgaggc ttg

23

<210> 381

<211> 21

<212> DNA

<213> Homo sapiens

<400> 381

gtccacgttt ccagaacacc c

21

<210> 382

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 382

cgaaatccca aagacacaga cc

22

&lt;210&gt; 383

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 383

gagttgctga agctgcggtg ag

22

&lt;210&gt; 384

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 384

gaaaaaggga gcttctgtgc atc

23

&lt;210&gt; 385

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 385

aacaagggga cagggactgg

20

&lt;210&gt; 386

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 386

cagaaacctg ggagcagatc c

21

&lt;210&gt; 387

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



<400> 387  
gcaacaagag gacatgcacc taag

24

<210> 388  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 388  
aaaggtagag gacatgccaa agc

23

<210> 389  
<211> 21  
<212> DNA  
<213> Homo sapiens

<400> 389  
ggagcagcca acaactcggt c

21

<210> 390  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 390  
gctcatttaa tccccacaac acc

23

<210> 391  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 391  
ccaccacacc tggctaattt ttg

23

<210> 392  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 392  
tgtagctcac cctctggact ttg

23

<210> 393

<211> 20

<212> DNA

<213> Homo sapiens

<400> 393

aatatgcaac cctcccctgc

20

<210> 394

<211> 23

<212> DNA

<213> Homo sapiens

<400> 394

tgaggcagaa gaatcgcttg aac

23

<210> 395

<211> 22

<212> DNA

<213> Homo sapiens

<400> 395

acttgtcatt ggctgtcccc tc

22

<210> 396

<211> 21

<212> DNA

<213> Homo sapiens

<400> 396

tttgctgca ctgtgctttt g

21

<210> 397

<211> 23

<212> DNA

<213> Homo sapiens

<400> 397

ccatactcag catcctgcac tcc

23

<210> 398

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 398

aacgacagca accaggtgg

20

&lt;210&gt; 399

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 399

cagcaggtgt atctaattggc agg

23

&lt;210&gt; 400

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 400

agagtgggtgg ggagatgagg tg

22

&lt;210&gt; 401

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 401

tgcaaagggg cagactagag g

21

&lt;210&gt; 402

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 402

cgaccacttt tatgggagga gc

22

&lt;210&gt; 403

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 403  
ccaggtgttc tgaaccacac ttc

23

<210> 404  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 404  
tgtgtgatgt cgaggcttgt acc

23

<210> 405  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 405  
gaatgcaggg aagagaaggc ag

22

<210> 406  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 406  
gccatcagga catggtgatt tc

22

<210> 407  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 407  
tgagcactcg tacttcccga ag

22

<210> 408  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 408  
tgagagcagc agggatgact tc

22

<210> 409

<211> 23

<212> DNA

<213> Homo sapiens

<400> 409

aaactcctga tagccactgg tgg

23

<210> 410

<211> 20

<212> DNA

<213> Homo sapiens

<400> 410

agatcctccg ccgaaatgtc

20

<210> 411

<211> 24

<212> DNA

<213> Homo sapiens

<400> 411

ttactcttct ctgagatgcc cgag

24

<210> 412

<211> 22

<212> DNA

<213> Homo sapiens

<400> 412

tgggcagtgg cattgtctac tc

22

<210> 413

<211> 20

<212> DNA

<213> Homo sapiens

<400> 413

ttccaggagg tggcatttcc

20

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/19094

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : Please See Extra Sheet. US CL : 530/350; 435/69.1, 71.1, 71.2, 325, 252.3, 254.11, 471, 6 According to International Patent Classification (IPC) or to both national classification and IPC														
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 530/350; 435/69.1, 71.1, 71.2, 325, 252.3, 254.11, 471, 6 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched NONE Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.														
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>														
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
A	US 5,599,905 A (MOSLEY et al.) 04 February 1997 (04.02.97), see entire document.	1-10, 12-20												
A	US 5,356,795 A (LEONARD et al.) 18 October 1994 (18.10.94), see entire document.	1-10, 12-20												
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.														
<table border="0"> <tr> <td>* Special categories of cited documents:</td> <td>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td> </tr> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td> </tr> <tr> <td>"B" earlier document published on or after the international filing date</td> <td>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td> </tr> <tr> <td>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td> <td>"&amp;" document member of the same patent family</td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td></td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> </tr> </table>			* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	"B" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family	"O" document referring to an oral disclosure, use, exhibition or other means		"P" document published prior to the international filing date but later than the priority date claimed	
* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention													
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone													
"B" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art													
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family													
"O" document referring to an oral disclosure, use, exhibition or other means														
"P" document published prior to the international filing date but later than the priority date claimed														
Date of the actual completion of the international search 24 OCTOBER 2000		Date of mailing of the international search report 14 NOV 2000												
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer <i>Jayle Bridges</i> PREMA MERTZ Telephone No. (703) 308-0196												

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/19094

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:  
1-10, 12-20

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/19094

### A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C12N 5/10, 15/12, 15/63, 15/64; C07K 14/47, 14/705, 14/715; G01N 33/53, 33/567

### B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, CAS ONLINE, MEDLINE, CAPLUS

search terms: IL4 receptor alpha polymorphic variant, polynucleotide, genotyping, oligonucleotide, method, screening, assay

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-10, 12, drawn to a polynucleotide encoding a polymorphic variant of IL4R alpha, a recombinant organism, the polymorphic variant of IL4R alpha and a method of screening for drugs targeting the polymorphic variant.

Group II, claim 11, drawn to an antibody to the variant.

Group III, claims 13-16, drawn to a composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in a IL4R alpha gene.

Group IV, claim 17-20, drawn to a method of genotyping IL4R alpha.

Group V, claim 21, drawn to a method of predicting a haplotype pair for the IL4R alpha.

Group VI, claims 22-24, drawn to a method of identifying an association between a trait and at least one genotype or haplotype of the IL4R alpha gene.

Group VII, claims 25-26, drawn to a computer system and a genome anthology for the IL4R alpha gene.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Pursuant to 37 C.F.R. § 1.475 (d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention (Group I) comprises the first-recited product, a polynucleotide encoding a polymorphic variant of IL4R alpha, a recombinant organism, the polymorphic variant of IL4R alpha and a method of screening for drugs targeting the polymorphic variant. Further pursuant to 37 C.F.R. § 1.475 (d), the ISA/US considers that any feature which the subsequently recited products and methods share with the main invention does not constitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.